بسم الله الرحمن الرحيم

بناء اختبار تكيفي لقياس القدرة الرياضية وفق الإستراتيجية ثنائية المرحلة في نظرية الاستجابة للفقرة

أعداد

صباح جميل فدعوس الشديفات

إشراف

الدكتور نضال كمال الشريفين

حقل التخصص – القياس والتقويم

2008 م

بناء اختبار تكيفي لقياس القدرة الرياضية وفق الإستراتيجية ثنائية المرحلة في نظرية الاستجابة للفقرة

إعداد

صباح جميل فدعوس الشديفات

بكالوريوس معلم صف، جامعة الزرقاء الأهلية، 2005

قدمت هذه الرسالة استكمالاً لمتطلبات الحصول على درجة الماجستير، تخصص قياس وتقويم

في جامعة اليرموك، اربد، الأردن.

وافق عليها

د.نضال كمال الشريفين..........................رئيساً

أستاذ مساعد في التقويم، جامعة اليرموك

أ.د.محمد مقبل عليمات.............................. عضواً

أستاذ في التربية المهنية ومناهج التدريس، جامعة اليرموك

د.أحمد يوسف قواسمة.. عضواً

أستاذ مشارك في القياس والتقويم، جامعة اليرموك

د.ساري سليم سواقد..................................... عضواً

أستاذ مشارك في القياس والتقويم ، جامعة اليرموك

تاريخ المناقشة 2008/4/28م

الإهداء

إلى من رسمت لي الدرب قبل وفاتها وتركت فراغاً لا يستطيع أحد ملئه إلى أمي رحمها الله .

إلى من قدم لي كل الدعم وذلل أمامي كل الصعاب إلى زوجي الغالي محمد شديفات أمده الله بالصحة والعافية .

إلى من قدم لي يد المساعدة والحنان العم أبو شاتي وعائلته أدامهم الله .

إلى من كان إنجاز هذا العمل على حساب وقتهم إلى أولادي الأعزاء حفظهم الله .

إلى منبع الدعاء الصادق أخواني (نبيل، محمد، مصطفى، أمنور) .

" أهدي جهدي المتواضع هذا "

الباحثة
صباح الشديفات

شكر وتقدير

بعد أن انتهت أيام وساعات العمل والبحث والمراجعة أعود بذاكرتي إلى الوراء فــأرى جمعاً من الأشخاص يعجز لساني عن شكرهم فقد ذللوا الصعوبات وجعلوا العسير يسيراً وأولهم أستاذي الذي منحني الكثير من وقته والغزير من علمه إلى أستاذي الدكتور (نضال الــشريفين) كل الشكر والتقدير .

وأتقدم بالشكر الجزيل إلى أساتذتي أعضاء لجنة المناقشة: الاستاذ الدكتور محمد عليمات والدكتور أحمد قواسمة ً والدكتور ساري سواقد .

على تكرمهم بالموافقة على مناقشة هذه الرسالة وعلى ما أبدوه من ملاحظات أثرت هذا البحث وزادت من قيمته العلمية، كما أتقدم بالشكر الجزيل إلى الحاج أحمد شحاده شديفات كــل الشكر على جهوده العظيمة كما أتقدم بالشكر الجزيل إلى المشرفين والمعلمات ومدراء المدارس التي شكلت عينة الدراسة، كما أشكر كل من قدم عوناً معنوياً لإنجاز هــذا العمــل وجــزى الله الجميع كل الخير.

الباحثة /صباح الشديفات

فهرس المحتويات

فهرس الجداول

فهرس الأشكال

فهرس الملاحق

الملخص

شديفات، صباح جميل. بناء اختبار تكيفي لقياس القدرة الرياضية وفق الإستراتيجية ثنائية المرحلة في نظرية الاستجابة للفقرة . رسالة ماجستير، جامعة اليرموك . 2008
(المشرف :د. نضال كمال الشريفين).

هدفت هذه الدراسة إلى بناء اختبار تكيفي لقياس القدرة الرياضية وفق الإستراتيجية ثنائية المرحلة في نظرية الاستجابة للفقرة . ولتحقيق هذا الهدف قامت الباحثة ببناء اختبار قدرة رياضية من نوع الاختيار من متعدد في مبحث الرياضيات وحدتي (الضرب والقسمة) للصف الثاني الأساسي مؤلف من (60)فقرة بصورته النهائية.

ظمَ الاختبار (3) عينات هي:

- عينة التجريب الأولي وبلغت (100) طالب وطالبة.
- عينة تدريج الاختبار وبلغت (500) طالب وطالبة.
- عينة القياس التكيفي وبلغت (100) طالب وطالبة.

وكانت هذه العينات مختارة من مدارس أساسية تتبع مديريات التربية والتعليم الثلاث في منطقة محافظة المفرق .

تم تطبيق الاختبار على عينة الدراسة وحللت النتائج باستخدام البرمجيات التالية (SPSS ،BIGSTEPS،Notpad) حيث تم الحصول على تقديرات لقدرة الأفراد وتقديرات لقيم صعوبة الفقرات و الأخطاء المعيارية في قياسها وإحصاءات الملاءمة لنموذج الاختبار وأظهرت النتائج فاعلية استخدام الاختبارات التكيفية لطلبة الصف الثاني الأساسي كذلك أظهرت النتائج مستوى مقبول من الدقة في تقدير قدرات الأفراد عند استخدام الاختبار التكيفي بدلاً من "الاختبار التقليدي".

الكلمات المفتاحية : نظرية استجابة الفقرة، الإستراتيجية الثنائية، اختبار تكيفي، قدرة الفرد، صعوبة الفقرات.

الفصل الأول

خلفية الدراسة وأهميتها

الفصل الأول
خلفية الدراسة وأهميتها

المقدمة

ساهم القياس النفسي والتربوي في تطور العلوم التربوية، حيث شمل ذلك التطور مجالات الاختبارات النفسية والتربوية المختلفة. ويمكن القول أن قياس تحصيل الطلبة وتقدمهم في المسارات العلمية كان ولا يزال جزءا مهما من الاهتمام التربوي، منذ أن كانت هناك علاقة بين معلم وطالب، ومن المحطات الهامة في تاريخ القياس النفسي والتربوي النظرية الكلاسيكية في القياس (Classical Test Theory) والتي ارتكزت إلى افتراض التوزيع السوي لدرجات المفحوصين على السمة موضوع القياس. وذلك بمقارنة درجات المفحوص بمعايير المجموعة التي ينتمون إليها، وعليه فالدرجة الكلية للمفحوص في الاختبار تعتمد على سهولة أو صعوبة فقرات الاختبار، ويعتمد تفسير تلك الدرجة على المجموعة التي ينتمي إليها المفحوص. فأدى ذلك إلى ظهور الاختبارات معيارية المرجع (Norm Referenced Test)، والتي تعتمد على قيمتي المتوسط والانحراف المعياري للمجموعة المعيارية. وبذلك يتغير تفسير نتائج القياس بتغير المجموعة المعيارية (كاظم،1998). ثم ظهرت الاختبارات محكية المرجع (Criterion Referenced Test) والتي يرجع فضل ظهورها إلى العالم جليزر Glazer (1963) المشار إليه في (علام، 2001). حيث يفسر أداء الفرد بناء على محتوى الاختبار ذاته، وليس على أداء الأفراد الآخرين كما في الاختبارات معيارية المرجع.

وظل العاملون بالقياس النفسي يستخدمون مبادئ وأسس النظرية التقليدية (الكلاسيكية) وبالاعتماد على مفاهيمها وافتراضاتها في بناء الاختبارات وتحليل نتائجها ولكن مع زيادة انتشار تطبيقاتها ظهرت جوانب من القصور في هذه النظرية وأصبحت تلك الجوانب أوجه قصور يشار إليها من قبل الباحثين في النظرية الكلاسيكية.

أوجه قصور في النظرية الكلاسيكية

- اعتماد معامل صعوبة الفقرة ومعامل تمييزها في تطوير الاختبار، على عينات المفحوصين الذين تم من خلالهم الحصول على هذه المعاملات، فمعاملات الصعوبة تكون أعلى عندما تكون عينة المفحوصين من ذوي القدرات العالية، و تكون متدنية إذا كانوا من ذوي القدرات المتدنية، وتكون معاملات التمييز مرتفعة إذا كانت العينة غير متجانسة في قدراتها، ومتدنية إذا كانت العينة متجانسة.
- تعتمد نتائج المفحوصين في النظرية الكلاسيكية، على عينة الفقرات المستخدمة في تقدير قدرة المفحوصين، الذين طبق عليهم الاختبار.
- تساوي تباين خطأ القياس بالنسبة لجميع المفحوصين هو الافتراض الذي تقوم عليه النظرية الكلاسيكية.
- جوانب قصور تعجز النظرية الكلاسيكية في تقديم حلول لها مثل:
- الكشف عن التحيز في الفقرات والمعادلة بين علامات الاختبارات المختلفة. (Hambelton & Swaminathan, 1985)
- هناك طرق في تقدير الثبات للاختبار في النظرية التقليدية منها الصور المتكافئة (Parallel Form) وهذه يصعب الحصول عليها عملياً. وحساب الثبات بالإعادة وهذه لها عيوب منها تعرض المفحوصين للنسيان.

وقد بذل علماء القياس جهودا كبيرة لتطوير نظرية حديثة في القياس النفسي والتربوي لمعالجة عيوب النظرية الكلاسيكية في القياس التي انتقدها كثيرا من العلماء مثل رايت (Wright) ، لورد (Lord)، هامبلتون (Hambelton). وقد نشر العالم لورد (Lord) عام (1952) كتابا جاء كنتيجة لجهود علماء القياس في تطوير نظرية حديثه في القياس النفسي

والتربوي لمعالجة عيوب النظرية الكلاسيكية في القياس وقد أطلق لورد على هذا الكتاب اسم "النظريات الإحصائية لدرجات الاختبارات العقلية" (Statistical Theories of Mental Test Scores) وبذلك وضع لورد أسس نظرية حديثة في القياس . وقد سميت هذه النظرية الحديثة بنظرية الاستجابة للفقرة (Item Response Theory). حيث تفترض هذه النظرية وجود دالة احتمالية بين معلمين (Parameters) أحدهما يتعلق بقدرة الفرد المفحوص والآخر يتعلق بالفقرة التي يُختبر بها. وبالتالي فإن هذه النظرية تهدف للتوصل إلى قيم تقديرية، لكل من هذين المعلمين ثم تستخدم هذه القيم في تقدير احتمال الاستجابة الصحيحة لكل فقرة من فقرات الاختبار (Linden, 1981).

وقد عرفت هذه النظرية كذلك بأسماء متعددة مثل" نظرية السمات الكامنة " Latint Trait Theory" أو" نظرية منحنى خصائص الفقرة " ICC" (Item characteristic curve) Theoryكذلك "نظرية الاستجابة للفقرةItem Response theory" وفي عام (1978) بدأت المؤسسات المدنية والعسكرية في الولايات المتحدة الأمريكية باستخدام هذه النظرية والبحث في المزيد من تطبيقاتها (Warm, 1978) .

وتتميز هذه النظرية بأنها تسمح بتقدير معالم فقرات الاختبار، بشكل مستقل عن عينة الأفراد الذين يستجيبون للاختبار (Person Free). كما يمكن تقدير قدرات الأفراد، بشكل مستقل عن عينة الفقرات التي تقدم لهم (Item Free). لذا بدأ علماء القياس المعاصر، في تطبيق أسس النظرية الحديثة في نظام بنوك الأسئلة (Item Banking) ونظام القياس التكيفي (Adaptive Testing)الذي يقوم على المواءمة بين صعوبة الفقرات وقدرة الأفراد، بحيث تكون صعوبة الفقرات مناسبة لمستوى قدرة الأفراد الذين يُمتحنون بها (Lord, 1980).

حيث استندت هذه الدراسة إلى مجموعة من المرتكزات، تشكل الإطار النظري لها، هي مستخلصة من نظرية الاستجابة للفقرة والقياس التكيفي وفق الإستراتيجية ثنائية المرحلة وفيما يلي أهم مرتكزات الإطار النظري لهذه الدراسة:

نظرية استجابة الفقرة (Item Response Theory)

لقد مرت نظرية استجابة الفقرة في القياس بمحطات هامة سبقت كتاب لورد فبدايات نظرية إستجابة الفقرة تعود إلى فيرجسون (Ferguson) وتلا ذلك مقالات لبروجدن (Brogden) وكذلك كرونباخ وورنغتون (Cronbach and Warington) كذلك يطلق على نظرية الاستجابة للفقرة الاختبارية اسم آخر، هو نظرية السمات الكامنة (Latent Trait Theory). حيث تفترض هذه النظرية وجود سمة لدى الفرد، يطلق عليها السمة الكامنه التي يُعزى لها الاتساق في أداء المفحوصين في عدة اختبارات. وبما أنه لا يوجد للسمة مدلول مادي ولا يمكن ملاحظتها مباشرة فيطلق عليها تعبير "السمة الكامنة" (Warm, 1978). وتستخدم هذه النظرية الرمز (θ) للدلالة على كمية السمة وتقرأ "ثيتا".

وتمتد (θ) على متصل يمتد من $-\infty$، إلى ∞ وقد بين هامبلتون وسوامنثان وروجرز (Hambelton. Swaminathan, Rogers, 1991) مزايا نظرية السمات الكامنة:

- تقدير قدرة الفرد بشكل مستقل عن عينة الفقرات المطبقة عليه(Item Free).
- تقدير خصائص سيكومترية للفقرات، بشكل مستقل عن عينة الأفراد التي استخدمت في تقدير هذه الخصائص (Person Free).
- قياس درجة الدقة في تقدير كل قدرة لكل فرد على حده، بواسطة الخطأ المعياري في التقدير.

- تعتمد احتمالية الاجابة الصحيحة للفرد على الفقرة على كل من قــدرة الفــرد، وصــعوبة الفقرات. وتتمثل بنقاط على متصل واحد مشترك.
- النموذج المستخدم، يربط بين فقرات الاختبار، ومستويات القدرة للأفراد. وقد يكون النموذج أحادي المعلم أو ثنائي المعلم أو ثلاثي المعلم.
- لا تلعب السرعة دور في الإجابة عن فقرات الاختبار.
- وقد أضاف هامبلتون وسوامينيثان (Hambleton & swaminarham, 1985) ميــزة أخــرى تتعلق بإحصائي يشير إلى الدقة في تقدير قدرة كل فرد وهو الخطــأ المعيــاري للتقــدير (Standard Error of Estimation) وهذا الإحصائي مختلف من فرد لآخر.

افتراضات نظرية السمات الكامنة.(Assumptions of Latent Trait Theory)

ترتكز نماذج السمات الكامنة على مجموعة من الافتراضات يجب توافرها في البيانــات المستمدة من الاختبار، وتختلف هذه النماذج في عدد الفروض اللازم توافرها في البيانات ويعــد نموذج (راش) أقلها جميعاً في عدد الفروض (علام ،1995) . ويذكر (هامبلتون وسواميناثان) عــدداً منهــا هــي: أحاديــة البعــد والاســتقلال الموضــعي، والمنحــى المميــز للفقــرة (Hambelton & Swaminathan,1985) وفيما يأتي توضيح لهذه الافتراضات :

أحادية البعد : Unidimensiomality

يعد فرضاً أساسياً في نماذج السمات الكامنة جميعها وهو المصطلح التقني لصياغة فكرة أن جميع الفقرات تقيس سمة كامنة واحدة . وتفسر انستازي ويوربينـــا (Anastasia & Urbina, 1997) أحادية البعد بأن الاستجابة على الاختبار تعزى إلى سمة مفردة (Single Trait) وعموماً فإن الفرض يتحقق بصورة مناسبة إذا كان الأداء على الاختبـــار يعتمـــد علـــى ســـمة واحـــدة (Predominat Trait) حتى عندما تؤثر سمات أخرى على الأداء بشكل ثـــانوي ومـــن الناحيـــة التطبيقية لا يوجد مقياس نفسي أحادي البعد تماماً، بالمقابل فإن هناك أبعاداً ثانوية عديدة تـــؤثر في الاستجابة على الفقرة، وتركز هذه النظرية على قضية أساسية هي ما إذا كانت الأداة أحادية البعد لدرجة كافية تسمح بتطبيق النظرية، إذ أن تطبيق النظرية يعتمد على هذا الافتراض لـــذا فمن المهم التحقق من مدى تحقيق البيانات لهذا الافتراض.

ويتحقق افتراض أحادية البعد عندما يكون النجاح على الفقرة أو الفشل في الإجابة عنها يعتمد على القدرة المقيسة، فقط ولا شيء غيرها، ويبين لورد (1980) أن الأداء علـــى الفقـــرة يعتمد جزئياً على سمة أخرى غير السمة (أو القدرة) المقيسة وذلك عندما تقدم المعرفـــة عـــن الإجابة على الفقرة معلومات أخرى عدا ما هو مستنبط من القدرة، وهذا يناقض فرض أحاديـــة البعد (Lord, 1980). وعلى ذلك فإن تحقيق أحادية البعد يتطلب وجود عامل مسيطر يؤثر على الأداء في الاختبار، وهو القدرة التي يقيسها الاختبار.

ويميز هامبلتون (Hambelton) بين مفهومي أحادية البعـــد والتفـــرد (Singularity)، إذ وضح أن أحادية البعـــد لهـــا إســـهام فـــي الاختبـــارات العقليـــة مـــع أنهـــا تتـــضمن عـــدداً من القدرات الفرعية، لذلك فإنها لا تعبر عن بناء مفرد، بل يمكن عدها مركباً من عدة عناصر، لـــذلك فمـــن الـــصعب بنـــاء اختبـــار مفـــرد (Singular) مـــن الناحيـــة العمليـــة

(Hambelton & Swamanthen, 1985) ويطرح هاتي (Hattie, 1985) مؤشرات خمسة تستخدم أدلة على أحادية البعد هي : نمط الاستجابة، ومعامل الثبات أو الاتساق المحسوب بمعامل ارتباط الفقرة مع الدرجة الكلية، والمكونات الرئيسية للأداة، والتحليل العاملي، ومطابقة البيانات لأحد نماذج السمات الكامنة ويمكن التحقق من هذه المؤشرات كما يأتي :

- **نمط الاستجابة (Indices based on Answer Patterns)**: وذلك باستخدام نموذج (جتمان) (Guttman) في الاستجابة على الفقرات، حيث يفترط جيتمان أن السمة تكون أحادية البعد إذا أمكن تحقيق مفهوم التراكمية بحيث إجابة سؤال في مستوى ما تعتمد على قدرة المفحوص على إجابة سؤال في مستوى أدنى . كذلك مؤشر التجانس (Index of Homogeneity) ومؤشر جريين (Green Index) .
- **الاتساق الداخلي:** يعد معامل ارتباط الفقرة بالدرجة الكلية مؤشراً على اتساق ما تقيسه هذه الفقرات وبالتالي مؤشراً على أن الفقرات تقيس سمة واحدة، ويشار في هذا الصدد إلى أن أهم إحصائي في عملية بناء الاختبار هو معامل ارتباط الفقرة بالدرجة الكلية فالفقرات التي تتمتع بمعامل ارتباط مرتفع يحتفظ بها في الاختبار، لأن ما تقيسه مثل هذه الفقرات ينسجم مع ما يقيسه الاختبار بشكل عام، وهذا يعطي مؤشراً على أن مجموعة الفقرات تقيس سمة أو قدرة واحدة (Lord, 1980) ومن أمثلة ذلك معامل ثبات كرونباج ألفا (Gronbach α) أو معامل كودر – ريتشاردسون (20) كحالة خاصة من كرونباخ α كذلك متوسط معاملات الارتباط الداخلية للفقرات (Inter- Item Correlation) .
- **المكونات الرئيسية : (Indices Based on Principal Components)** وهي مؤشرات تعتمد على التحليل العاملي (Factor Analysis) في تحقيق افتراض أحادية البعد للاختبار.

- **التحليل العاملي (Factor Analysis):** أقترح(لورد)دلالة إحصائية تقريبية في حساب الجذور الكامنة (Eigen Values) في التحليل العاملي لتحقيق أحادية البعد لمجموعة فقرات الاختبار، وهي كون قيمة الجذر الأول كبيرة مقارنة بالجذور الكامنة الأخرى التي تقترب قيمتها من بعضها بعضا. ويضيف (لورد) توضيحا لما سبق في أن أحادية البعد أكثر عمومية من التحليل العاملي الذي يشترط التوزيع الاعتدالي لدرجة المفحوصين، في حين أن أحادية البعد هي خاصية للفقرات ولا تعتمد على توزيع القدرة لمجموع المفحوصين (Lord , 1980).

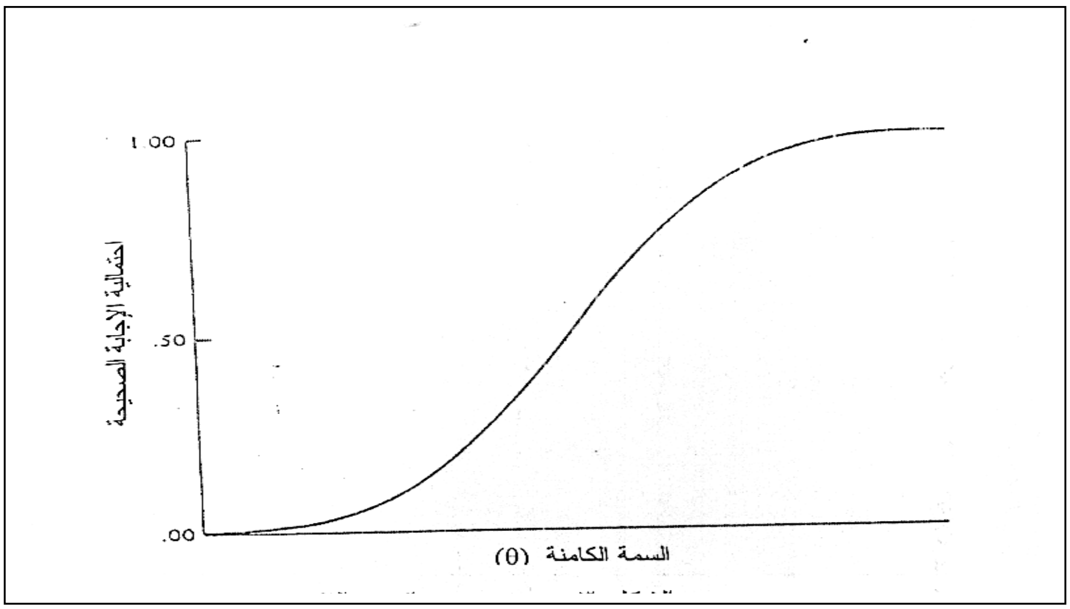
- **مطابقة البيانات لأحد نماذج السمات الكامنة:** (Indices Based on Latent Trait Models) يمكن اعتماد الاستقلال الموضعي (Local Independence)لكل زوج من الفقرات الاختبارية ولأي مجتمع متجانس من المفحوصين كمؤشر لأحادية البعد، إذ يترتب على الاستقلال الموضعي وجود سمة واحدة يقيسها الاختبار (Crocker & Algina, 1986).

الاستقلال الموضعي :Local Independence

هذا الافتراض مكافئ لافتراض أحادية البعد ويعني أن استجابات المفحوصين على الفقرات المختلفة في الاختبار مستقلة إحصائياً، أي أن أداء المفحوص على فقرة معينة يجب أن لا تؤثر سلباً أو إيجاباً على أدائه على أي فقرة أخرى في الاختبار، أي لا يتضمن محتوى أي فقرة أي إشارة أو توضيح لإجابة الفقرات الأخرى في الاختبار. وعندما يتحقق افتراض الاستقلال الموضعي لفقرات الاختبار فإن احتمالية إجابة كل مفحوص على أي زوج من الفقرات هو حاصل ضرب احتمال إجابة المفحوص على كل فقرة منها على حدة.

منحنى خصائص الفقرة Item- Characteristic Curve, ICC

يمثل منحنى خصائص الفقرة (ICC) أحد المفاهيم الأساسية لنظرية الاستجابة للفقــرة، ويمثل هذا المنحنى العلاقة بين احتمالية الإجابة الصحيحة على الفقرة والسمة الكامنة (θ)، وفي معظم تطبيقات نظرية الاستجابة للفقرة يأخذ هذا المنحنى شكل (S)، وكما هو موضح في الشكل (1) فإن زيادة قيمة السمة الكامنة (θ) يتبعه زيادة في احتمال الإجابة الصحيحة عــن الفقــرة ويقترب خط الاستواء السفلي (Lower Asymptote) فيه من الصفر فيما يقترب خط الاســتواء العلوي (Upper Asymptote) من (1).



الشكل (1): شكل توضيح لمنحنى خصائص الفقرة

* **المصدر ((Crooker and Algina , 1986P(34)**

ويفسر هذا المنحنى وجود مجتمعات فرعية من المفحوصين الذين يمتلكون نفس قيمــة السمة الكامنة وبالتالي تتساوى لديهم احتمالية الإجابة الصحيحة علــى الفقــرة، أمــا الأفــراد

الذين هم بمستوى قدرة أعلى فان احتمال إجابتهم على الفقرة إجابـــة صـــحيحة تكـــون أعلـــى. (Crocker and Algina , 1986).

التحرر من السرعة في الأداء Speededness :

أن الافتراض الضمني عند استخدام جميع نماذج الاستجابة للفقرة هـــو أن الاختبـــارات التي ستطابق النماذج لم تطبق تحت شرط السرعة. أي أن المفحوصين الذين يفشلون في الإجابة عن فقرات الاختبار يرجع ذلك إلى محدودية قدرتهم وليس بسبب عدم تمكنهم محاولة حل جميع الفقرات ضمن الاختبار وهذا الافتراض متضمن في افتراض أحادية البعد لأنـــه عنـــدما تـــؤثر السرعة على الأداء في الاختبار فان هناك سمتين تؤثران على الأداء في الاختبار همـــا ســـمة السرعة في الأداء، والسمة التي تقاس بمحتوى الاختبار، ويمكن تقييم هذا الافتراض من خـــلال حصر عدد الأفراد الذين فشلوا في إنهاء مجموعة فقرات الاختبار المطبقة ضمن الوقت المحدد للاختبار.

نماذج النظرية الحديثة

لقد انبثقت عن نظرية القياس الحديثة مجموعـــة مـــن النمـــاذج كمـــا عرفـــت كـــذلك بنماذج السمات الكامنة (Latent Trait Models) ويحتوي كل منها على معادلة رياضـــية تحـــدد علاقة أداء الفرد على فقرة من فقرات الاختبار بقدرته التي تكمـــن وراء هـــذا الأداء وتفـــسره (Hambelton & Swaminathan, 1985) وهذه النماذج منها الأحادي وهو النموذج الذي يفترض وجود قدرة واحدة تفسر أداء الفرد في الاختبار ومنها نماذج متعددة الأبعاد وهي التي تفتـــرض وجود اكثر من قـــدرة واحـــدة تفـــسر هـــذا الأداء. وقـــد أشـــار هـــامبلتون وســـوا مينثـــان (Hamblton & Swaminathan, 1985) إلى أشهر طرق تصنيف النماذج أحادية البعـــد وهـــو تصنيفها تبعاً لاستجابات المفحوصين ومستويات الاستجابات الـــشائعة هـــي:ثنائيـــة التـــدريج

(Dichotomous) ومتعددة التدريج (Polytomous)، واستجابات متـصلة (Continuous) وهـذا الاختلاف في الصور الرياضية لهذه النماذج أدى إلى ظهور العديد من هذه النماذج كان أكثرها شيوعاً واستخداماً في مجال القياس النفسي والتربوي، ونموذج الاوجايف التراكمـي الطبيعـي (Cumulative Normal Ogive) والنماذج اللوغارتمية (Logistic Models).

تصف هذه النماذج العلاقة بين السمة الكامنة موضوع القياس واحتمال الإجابة الصحيحة عن الفقرة وهذه العلاقة يمكن وصفها بثلاث معلمات وهي: معلمة الصعوبة، ومعلمة التمييـز، ومعلمة التخمين، وتختلف هذه النماذج من حيث الافتراضات التي تقوم عليها والاقتران الذي يتم بواسطته حساب المعلمات **وهذه النماذج هي :**

1- **النموذج اللوجستي الأحادي المعلم (نمـوذج راش) (One-Parameter Logistic Model (Rasch Model)**

يرتبط هذا النموذج باسم عالم الرياضيات الدنماركي جورج راش عام (1966) ويستخدم هذا النموذج لتقدير صعوبة الفقرات ويفترض بان جميع الفقرات لها قيمة تمييزيـة واحـدة وان الفرد لا يلجأ إلى التخمين العشوائي في إجابته عن فقرات الاختبار. ويمكن التعبير عن نمـوذج راش بالمعادلة التالية كما أشار (Lord, 1980)

$$P_i(\theta) = \frac{e^{D(\theta - bi)}}{1+e^{D(\theta - bi)}}$$

حيث ((Pi(θ)) هي احتمال أن يجيب شخص قدرته (θ) على الفقرة (i) إجابة صحيحة، و(bi) = صعوبة الفقرة، و (D) = ثابت قيمته (1.7).

2- النموذج اللوجستي الثنائي المعلمات (Two-Parameter Logistic Model)

اقترح بيرنبوم (Birnbaum)عام 1962 النموذج اللوجستي ثنائي المعلم، وهذا النمــوذج يسمح بأن تختلف فقرات الاختبار في صعوبتها، وتمييزها ويمكن توضيح هذا النموذج بالمعادلة الآتية كما أشار لورد (Lord, 1980).

$$Pi(\theta) = \frac{e^{Dai\ (\theta - bi)}}{1 + e^{Dai\ (\theta - bi)}}$$

حيث :

(D) : ثابت قيمة (1,7) .

(ai): تمييز الفقرة .

(bi) : صعوبة الفقرة

($Pi(\theta)$): احتمال الإجابة الصحيحة لفقرة (i) للفرد الذي مستوى قدرته (θ)

(e): الأساس اللوغارتيمي الطبيعي = 2.7183

يفترض هذا النموذج أن الطلبة ذوي القدرة المنخفضة لا يمكنهم تخمين الإجابـــة عـــن الفقرات الصعبة إجابة صحيحة وبذلك يفترض النموذج أن معلم التخمين يـــساوي (صـــفر) وأن منحنيات خصائص الفقرة تختلف في ميلها حيث أن بعض المنحنيات تتزايد بشكل أســـرع مـــن الأخرى وبالتالي الفقرات المناظرة لهذه المنحنيات تميز بشكل أكبر من المنحيات الأخرى، كمـــا وتختلف في موقعها على متصل السمة أي أن بعض الفقرات أكثر صعوبة من الأخرى وبالتالي يسمح هذا النموذج باختلاف معلمي التمييز والصعوبة.

3- النموذج اللوجستي الثلاثي المعلمات (Three-Parameter Logistic Model)

يتكون هذا النموذج من ثلاثة معالم وهي (الصعوبة والتمييز والتخمين)فالمعلمة المضافة له عن المعلمات في النموذج الثنائي هي (C) و هي معلمة التخمين أي احتمال توصل الأفــراد ذوي القدرة المنخفضة إلى الإجابة الصحيحة عن الفقرة بمحض الصدفة أو التخمين.

ويعبر عنه عادة بالمعادلة الآتية التي وضحها لورد(Lord ,1980).

$$Pi(\theta)=Ci+(1-Ci)=\frac{e^{Dai\ (\theta-bi\ )}}{1+e^{Dai\ (\theta-bi\ )}}$$

حيث أن

(i=1,2,…..n) عدد الفقرات

Ci : معلمة التخمين

: معلمة صعوبة الفقرة.

ai : معلمة تمييز الفقرة

D :عامل ضبط التدريج(1,7)

يعتبر هذا النموذج أكثر النماذج عمومية لأنه يسمح باختلاف معالم الفقرة الثلاثة فيمكن أن تختلف هذه المنحنيات في ميلها (معلم التمييز)، وموقعها على متصل السمة (معلم الصعوبة)، وخط التقارب السفلي لمنحنى خصائص الفقرة (معلم التخمين).

أما الدالة الرياضية التي يعبر عنها النموذج الأحادي المعلمة فتربط بين احتمال إجابــة الفرد إجابة صحيحة على الفقرة ومعلم صعوبة الفقرة. ومنحنيات خصائص الفقرة لهذا النموذج تختلف فقط في موقعها على متصل السمة وبالتالي الفقرات تختلف في صــعوبتها ويعــد هــذا

النموذج حالة خاصة من النموذج الثلاثي المعلمة لذلك يمتلك هذا النموذج خصائص تجعل لـــه خصوصية لاستخدامه ومنها انه يتضمن أقل عدد من معالم الفقرة وهذا يسهل التعامل معه لأن مشكلة تقدير هذه المعالم أقل إلى حد كبير من النماذج الأكثر عمومية.

كما أضافتا ألين وين (Allen & yen, 1979) أن هناك ميزات عملية لنموذج راش تميزه عن نموذج المعلمتين والثلاثة معالم منها سهولة مطابقة البيانات للنموذج المستخدم، كما أنـــه لا حاجة لاشتقاق أوزان الفقرات للحصول على مجموع العلامة النهائية لإعطاء أكبر قـــدر مـــن المعلومات حول السمة الكامنة بالإضافة إلى سهولة التعامل مع هذا النموذج، ويتميـــز نمـــوذج راش بإمكانية تقدير قدرات الأفراد بطريقة مستقلة عن صعوبة الفقرات والتفاعل بـــين قـــدرة الفرد وصعوبة الفقرة هو الذي يحدد احتمال توصل الفرد إلى الإجابة الـــصحيحة عـــن الفقـــرة (علام1995)، وأوضح أن نموذج راش هو ميزان لوعاريتمي وحدة القياس المستخدمة لقيـــاس قدرة الأفراد وصعوبة الفقرات هي وحدة الترجيح اللوغاريتمي (Logit) وبالتالي فإن قدرة الفرد هي اللوغاريتم الطبيعي لترجيح الإجابات الصحيحة على الفقرات من النوع الذي استخدام فـــي تعريف نقطة صفر ميزان صعوبة الفقرات وصعوبة الفقرة هي اللوغاريتم الطبيعـــي لتـــرجيح الإجابة الخطأ على هذه الفقرات للأفراد الذين تقع قدراتهم عند نقطة صفر ميزان القدرة.

أن صعوبة الفقرة تمثل بنقطة على مقياس القدرة يكون عندها مقدار احتمـــال الإجابـــة الصحيحة عن الفقرة يساوي (50%) وهذه النقطة تناظر الإحداثي الأفقي لنقطة انقلاب المنحنى أي النقطة التي تكون عندها قدرة الفرد مساوية لصعوبة الفقرة وهذا يحدث عادة عندما تكـــون صعوبة الفقرة ملائمة تماماً لمستوى قدرة الفرد، ومعامل صعوبة الفقرة معرف نظريـــاً علـــى الفترة (- ∞، +∞) ويتراوح بين (2-،2+) وحدة معيارية، فالفقرات الـــسهلة، يقتـــرب معامـــل صعوبتها من (2+) وحدة معيارية، والفقرات متوسطة الصعوبة يكون معامل صعوبتها صـــفراً

أما معامل تمييز الفقرة فيمثل درجة العلاقة بين الأداء على الفقرة وقدرة الفرد وتتناسب قيمتـــه مع ميل منحنى احتمال الإجابة على الفقرة عند نقطة تساوي قدرة الفرد وصعوبة الفقرة. وكمـــا هو الحال في معامل الصعوبة فإن معامل التمييز معرف نظريا على الفترة (−∞، + ∞) ويندر عملياً الحصول على قيم معاملات تمييز أكبر من (2+) وحدة معيارية، ويزداد ميـــل المنحنـــى المميز للفقرة بزيادة قيمة معامل التمييز (Hambleton & Swaminathan 1985).

بعد اختيار أحد نماذج الاستجابة للفقرة لمجموعة معينة من البيانات يجب تقدير معـــالم الفقرات والأفراد ومن اشهر أساليب التقدير تلك المعتمدة على دالة الأرجحية فقد أوضح هلـــن، دارسقو وبارسون (Hulin, Drasgow , & Parson, 1983) الفكرة الرئيسة التي تقوم عليها دالـــة الأرجحية لتقدير معلمة معينة أنها ببساطة اختيار للقيمة التي تجعل البيانات الملاحظة مطابقـــة بشكل كبير للنموذج المختار، فإذا افترضنا فقرة ثنائية التدريج بحيث تأخذ العلامة (1) في حال الإجابة الصحيحة وتأخذ العلامة (Zero) في حال الإجابة الخطأ، فطريقة الأرجحية تعتمد فـــي تقدير قيمة القدرة (θ) للمفحوص على نمط الإجابة والتي تأخذ قيمتين(0,1) فإذا افترضنا فقـــرة ثنائية التدريج حاول الإجابة عنها (25) مفحوص (17) منهم إجابتهم صـــحيحة و(8) إجـــابتهم خطأ بحيث أن إجابة كل مفحوص على الفقرة مستقلة عن إجابات الآخرين فإن دالة الأرجحيـــة (L) هي حاصل ضرب احتمالات إجابات المفحوصين كالتالي:

$$L=P^{17}Q^{8}$$

حيث أن :

P : احتمال الإجابة إجابة صحيحة عن الفقرة (صعوبة الفقرة).

Q : احتمال الإجابة إجابة خاطئة على الفقرة.

وبالتالي فإن قيمة (P) التي تعظم هذه الدالة وتجعل قيمتها أكبر ما يمكن هـــي التقــدير لصعوبة الفقرة (Hulin , Drasgow & Parson , 1983).

لقــد اورد لنكــر ورايــت (Lincare & wright, 1993) أن البرنــامج المــستخدم (BIGSTEPS) يستخدم طريقة (PROX) ثم طريقــة الأرجحيــة القــصوى غيــر المــشروطة (Unconditional Maximum Likelihood) (UNCON) والتي تسمى أيضاً طريقــة الأرجحيــة القصوى المشتركة (Joint Maximum Likelihood) (JML) بصفة تتابعيه لتقدير معالم القدرات للأفراد ومعالم الصعوبة للفقرات وفيما يلي توضيحاً لكل منهما:

طريقة (PROX) : وتهدف هذه الطريقة إلى تحويل استجابات أفراد العينة عن فقرات الاختبار إلى قيم تقديرية متحررة لكل من صعوبة الفقرات وقدرات الأفراد وذلك وفقاً للخطوات التالية:

1. إعداد مصفوفة البيانات الخاصة بفقرات الاختبار والتي تبين توزيع الدرجات التي حصل عليها كل فرد من أفراد العينة على كل فقرة وهي إما واحد (إذا كانت الإجابة صــحيحة) أو صفر (إذا كانت الإجابة خاطئة).
2. استبعاد الأفراد الذين أجابوا عن جميع الفقرات إجابة صحيحة والذين أجابوا عن جميــع الفقرات إجابة خاطئة.
3. تقدير قيم مبدئية لكل من صعوبة الفقرات وقدرات الأفراد.
4. حساب عاملي تعديل أحدهما للفقرات والأخر للأفراد للحصول على تقــديرات نهائيــة متحررة لكل من معالم صعوبة الفقرات وقدرات الأفراد وحساب الأخطاء المعياريــة لهــذه التقديرات (علام، 1995).

5. **طريقة الأرجحية القصوى المشتركة (JML):** وتستخدم بصفة تتابعية لتقدير أدق التقديرات لمعالم صعوبة الفقرات ومعالم قدرات الأفراد وذلك وفقاً للخطوات التالية المشار إليها في(الشريفين،2003).

1. تحديد تقديرات مبدئية لمعالم القدرة للأفراد ومعالم الصعوبة للفقرات.
2. جعل مركز تقديرات معالم الصعوبة للفقرات مساوياً لمتوسط صعوبات الفقرات والذي يساوي صفر.
3. أجراء عمليات إعادة متعاقبة (lterations) للحصول على مجموعة معدلة من معالم الصعوبة.
4. باستخدام المجموعة المعدلة من معالم الصعوبة نحصل على مجموعة معدلة من معالم القدرة بعد عمليات إعادة متعاقبة (lterations).
5. تعاد الخطوات (4,3,2) للحصول على التقديرات النهائية المتحررة لمعالم القدرة.

القياس التكيفي (Adaptive Testing)

من التطبيقات التي أصبحت ممكنة بعد ظهور نظرية الاستجابة للفقرة الاختبارات التكيفية، تلك الاختبارات التي توائم بين صعوبة الفقرات وقدرات الأفراد، بحيث تكون صعوبة الفقرات مناسبة لمستوى قدرات الأفراد الذين يمتحنون بها (Lord, 1980) بعكس الاختبارات التقليدية والتي يتم من خلالها تعريض كل المفحوصين لجميع الفقرات الواردة في الاختبار بغض النظر عن مستوى قدرة المفحوص.

وعند تطبيق القياس التكيفي يتم تقديم فقرات تختلف في الصعوبة من مفحوص لآخر اعتمادا على قدرة المفحوص، وذلك لأنه لا فائدة من تقديم فقرات ذات صعوبة متدنية لمفحوص قدراته مرتفعة، وكذلك لا فائدة من تقديم فقرات ذات صعوبة مرتفعة للمفحوص ذي القدرة

المتدنية علاوة على أن اختبار الأفراد بفقرات تناسب قدراتهم يؤدي إلى قياس فعال بحيث يـتم الحصول على اكبر قدر من المعلومات عن المفحوصين(Lincare , 2000).

ويطلق على الاختبارات التكيفية أحيانا أسماء أخرى منها الاختبارات المفصلة (Tailored Tests) أو اختبارات الفقرات المتسلسلة (Sequential Item Testing) أو المتفرعـة (Pranched) أو التفريدية (Individualized) أوالمبرمجة (Programmed) أو الدينامية (Dynamic) أو الإجابة المـشروطة (Response Contingent) ويمكـن تـصميمها بالورقـة والقلـم أو بالحاسـوب (Anastasia&Urbina, 1997).

وقد كانت اغلب الاختبارات التكيفية بداية تبنى باستخدام فقـرات ذات إجابـة ثنائيـة التدريج (dichotomous) والتي تأخذ بعين الاعتبار فيما إذا كان المفحوص قد عـرف الإجابـة الصحيحة أو خمنها أو اختار بديلا خطأ، إلا أن الكثير من الدراسات أشارت إلى فائدة استخدام الفقرات ذات الإجابة متعددة التدريج (Polytomous) في الاختبارات التكيفية (Dodd & Koch, 1992) وعموما فـإن نمـوذج التقـدير الجزئـي (Partial Credit Model) لـراش (Rasch) للاستجابات المتعددة التدريج والذي طوره فيما بعد ماسـترز (Masters) قـد وفـر المجـال لتطبيقات القياس التكيفـي الـذي يتـضمن الجوانـب المعرفيـة، والشخـصية والاتجاهـات (Koch& Dodd, 1989).

وقد أورد كنجزبري وزارا (Kingsbury & Zara, 1989) ست خطوات لبناء الاختبـار التكيفي هي :

1-بناء تجمع الفقرات (Item Pool)

حيث تعد الفقرات ذات المعالم المقدرة جيدا حسب نظريـة الاسـتجابة للفقـرة أمـراً ضروريا للاختبار التكيفي، ولا يوجد حجم معين بالنسبة لعدد فقرات التجمع، إلا أن التجمع الذي

يحتوي 100 فقرة يمكن أن يُعطي نتائج دقيقة، وأهم المعالم الموجودة في هذا التجمع هو معلـــم الصعوبة الذي ينبغي أن يكون متدرجاً.

2- اعتماد نموذج لاستجابة الفقرة

وينبغي اختيار النموذج في سياق نظرية استجابة الفقرة مثل النموذج الأحادي المعلم، أو الثنائي المعلم، أو الثلاثي المعلم، أو نموذج التقدير الجزئي في حال تكوّن الاختبار من الفقـــرات متعددة التدريج.

3- إعتماد الطريقة التي يتم فيها إختيار فقرات الإختبار

وهي الطريقة التي يتم فيها تحديد الفقرات التي تعطى للمفحوص في المرة الأولى ويمكن البدء بفقرات ذات مستويات متعددة من الصعوبة في تجمع الفقرات، ويستحب أن تكـــون تلـــك الفقرات المدخلية تغطي المناطق المختلفة للمحتوى وتراعي التوازن في التمييز، ويضاف إلـــى ذلك إمكانية البدء حسب اختيار المفحوص فـــي مـــا يـــسمى بالاختبـــارات التكيفيـــة الذاتيـــة (Self Adaptive Testing) .كذلك في الاستراتيجية ثنائية المرحلة تسمى الفقرات التـــي تعطـــى للمفحوص في المرة الأولى بالاختبار الاستطلاعي (Routing Test) وتجرى على أساس القـــدرة المقدرة بواسطته توجيه المفحوص لاختبار آخر يناسب قدراته من حيث صعوبة فقراته.

4-اعتماد طريقة اختيار الفقرات

وترتكز طريقة اختيار الفقرات إلى الإستراتيجية المستخدمة في الاختبار التكيفي ســـواء كان ثنائي المرحلة أو متعدد المراحل أو محوسباً، ويتفرع عن ذلك معايير اختيار الفقرات حسب معلـــم الـــصعوبة أو دالـــة المعلومـــات (Test Information) أو الطريقـــة البايبزيـــة (Bayesian Method).

5- اعتماد طريقة رصد العلامة

ويقصد بها اعتماد طريقة لتقدير القدرة لمجموعة إجابات المفحوص أي اعتماد طريقـــة الأرجحية القصوى (Maximum Likelihood) أو الطريقة البابيزية (Bayesian Method).

6-اعتماد قاعدة التوقف عن تقديم الفقرات للمفحوص .

يعتبر معيار التوقف في القياس التكيفي عنصراً هاماً من عناصر تقدير القدرة للمفحوص فعندما يكون الاختبار قصيراً جداً فإن تقدير القدرة يحتاج إلى تقديم جميع الفقرات وإلا فإن تقدير القدرة قد يكون غير دقيق، أما عندما يكون الاختبار طويلاً فأن الاختبار التكيفي يفقد مبرراتـــه من حيث الاقتصاد في الجهد والوقت وعدد الفقرات وهناك مجموعة أسس يتم إنهـــاء الاختبـــار بناءً عليها:

1. عندما تصل دقة التقدير إلى قيمة معينة حيث يتم التوقف عندها وتحدد هذه الدقة بقيمـــة الخطأ المعياري عند تلك القدرة. عندما لا يصل تقدير القدرة للمفحوص إلى انحـــرافيين معياريين بعيداً عن علامة القطع أو معيار الفشل والنجاح في الاختبـــارات المعياريـــة والذي يكون مقررا سلفاً أو بمعنى أخر عدم قدرة المفحوص على تغيير مستوى قدرتـــه مهما قدم له من فقرات الاختبار المتبقية.
2. عندما تكون إجابة المفحوص في الاختبارات المحوسبة سريعة جداً أو بطيئة جداً فـــإن البرنامج يتوقف عن تقديم أي فقرات للمفحوص بحيث يدل ذلك على عدم جديته.
3. هناك حد أدنى من الفقرات يجب أن تقدم قبل أن يتوقف الاختبار بحيـــث تغطـــي هـــذه الفقرات جميع جوانب المجال السلوكي المراد قياسه.

استراتيجيات القياس التكيفي :

تختلف الاختبارات التكيفية باختلاف الاستراتيجيات التي تتبعها تلك الاختبارات وقد قسم لورد (Lord, 1980) هذه الاستراتيجيات إلى ثلاثة استراتيجيات:-

أولاً : استراتيجية القياس التكيفي ذو المرحلتين (Two - Stage Strategy) :

وتسمى باستراتيجية القياس التكيفي ذي التفرع الثابت (Fixed Branch)وتتكــون هــذه الاستراتيجية من مرحلتين، المرحلة الأولى عبارة عن اختبار استطلاعي (Routing Test) يجري على أساس القدرة المقدرة بواسطته توجيه المفحوص إلى اختبار فرعي آخــر يناســب قــدرة المفحوص من حيث صعوبة فقراته (Lord, 1980).

وقد اقترح (Weiss,1983) استعمال فقرات الاختيار من متعدد باستخدام الورقة والقلم كتطبيق على هذه الإستراتيجية ،حيث يجيب كل مفحوص عن اختبار توجيهي أو (اســتطلاعي) قصير يصحح فوراً، وتستخدم نتائجه في توجيه كل مفحوص إلى واحد من اختبارات المرحلــة الثانية والتي عادة ما تكون أطول. والفكرة الأساسية هنا أن المفحوص ذو القدرة المتدنية يكون أداؤه ضعيفاً في الاختبار الاستطلاعي فيوجه إلى اختبار مرحلة ثانية أسهل لقياسه أما المفحوص ذو القدرة العالية تكون درجته أعلى في الاختبار الاستطلاعي ويوجه إلى اختبار مرحلة ثانيــة أكثر صعوبة لقياسه.

والإجراء العملي لاختيار فقرات الاختبار الاستطلاعي هو اختيار فقراته من بين فقرات الاختبارات الفرعية الخاصة بالمرحلة الثانية، فبعد تدريج تجمع الفقرات وتقسيمها إلى اختبارات فرعية ذات صعوبة متدرجة مختلفة، يتم انتقاء عدد قليــل مــن الفقــرات لتكــوين الاختبــار الاستطلاعي الذي يتم بواسطته تحديد مستوى قدرة ابتدائي للمفحوص الأمر الذي يترتب عليــه

توجيه المفحوص لأحد الاختبارات الفرعية الخاصة بالمرحلة الثانية (Ghiselli, Campell & Zedeck, 1981).

والميزة الرئيسية لهذه الطريقة إنها تزاوج بين مستوى صعوبة اختبار المرحلة الثانية ومستوى قدرة المفحوص، فلما كانت الاختبارات التقليدية تناسب الفئة المتوسطة من المفحوصين فأن طريقة الإستراتيجية ثنائية المرحلة تمتاز عنها بشكل أساسي للمفحوصين الذين يقعون على طرفي متصل القدرة أو السمة المقيسة (Lord, 1980) وقد كانت هذه الإستراتيجية هي المستخدمة في هذه الدراسة أثناء بناء اختبار تكيفي لقياس القدرة الرياضية في نظرية الاستجابة للفقرة الاختبارية. وقد تعرضت هذه الاستراتيجية لكثير من البحوث و الدراسات بهدف التطوير، ومن أمثلة هذا ما طرحه ماركو حيث يختار المفحوص أحد مستويات الاختبارات الفرعية بناء على ترتيب المستويات تنازليا حسب الصعوبة دون المرور بالاختبار الاستطلاعي ثم يتقدم لهذا الاختبارلقد قارن لورد (Lord, 1980) بين الاختبار ثنائي المرحلة والاختبار التقليدي بالنقاط التالية:

- يكون الاختبار التكيفي ثنائي المرحلة أفضل عبر مستويات القدرة جميعها عندما يكون الاختبار التقليدي ذات مدى واسع من الصعوبة لفقراته.
- عندما يكون مدى الصعوبة ضيقاً فإن الاختبار التقليدي يصبح أفضل لدى المفحوصين ذوي القدرة المتوسطة بينما يكون الاختبار التكيفي ثنائي المرحلة أفضل عند طرفي متصل القدرة.
- يكون الاختبار ثنائي المرحلة ذا قيمة أعلى عندما يمتد توزيع قدرات المفحوصين على مدى واسع من متصل القدرة أكثر مما يمكن قياسه بالاختبار التقليدي.

ثانياً : استراتيجية القياس التكيفي متعدد المراحل: (Multi_ Stage Strategy)

يعطى في هذه الإستراتيجية، فقرةمتوسطة الصعوبة للمفحوص؛ فإن أجابها إجابة صحيحة تقدم له فقرة تالية أكثر صعوبة، أما إذا كانت إجابته خاطئة، فتقدم له فقرة تالية أسهل وتعاد هذه الطريقة بعد كل استجابة حتى يعطي المفحوص عدداً من الاستجابات الصائبة أو الخاطئة. ويلزم لهذه الإستراتيجية تجمع اختباري كبير نسبياً (Anastasia & Urbana, 1997).

ثالثاً : استراتيجية القياس التكيفي المحوسب (Computerized Adaptive Testing Strategy)

وتتطلب هذه الإستراتيجية أن يتم اختيار الفقرات وحساب القدرة بشكل فوري باستخدام الحاسوب، فعادةً ما يبدأ الحاسوب بإعطاء فقرة ذات صعوبة متوسطة من تجمع الفقرات الموجودة، وبناءً على استجابة المفحوص يتم حساب تقدير لقدرته، ثم يتم تقديم الفقرة التالية والتي تعتمد على صحة أو خطأ إجابته السابقة، وتستمر هذه العملية حتى الوصول لما يسمى بقاعدة التوقف (Warm,1978)، وبالنسبة للفقرة التالية فإنها تكون ذات صعوبة أعلى من سابقتها إذا كانت الإجابة السابقة صحيحة، وذات صعوبة اقل من سابقتها إذا كانت الإجابة خطأ (Linacare , 2000) وعادة ما يتم تصميم قاعدة التوقف حسب رغبة الفاحص.

وقد أورد وورم (Warm, 1978) بعضا من قواعد التوقف وهي:

- التوقف عند استنفاد الفقرات الموجودة في تجمع فقرات بنك الأسئلة .
- التوقف عندما تصبح قيمة الخطأ المعياري في التقدير اقل من (0.0625) بينما يشيرامبرستون وريس (Emberston & Reise, 2000) إلى أن القيمة تلك يحددها الفاحص، لكن ذلك يتطلب حجماً كبيراً من الفقرات يتناسب مع القيمة المحددة، ويُعتبر هذا المعيار أفضل معيار للتوقف.

- التوقف عندما تصبح الفقرات الموجودة غير كافية .

وقد أضاف لينكير (Lincare,2000) بعض القواعد الأخرى مثل كون قيـــاس القـــدرة بعيداً جداً عن المحك، وان يُظهر المفحوص سلوكا يؤدي إلى إنهاء الاختبـــار مثـــل أن يكـــون المفحوص سريعا جدا في الإجابة أو بطيئاً جداً في الإجابة عن فقرات الاختبار، إلا أنه وبـــشكل عام لا يتوقف الاختبار قبل إن يتم تقديم الحد الأدنى من الفقرات، وتغطية جميع الفقرات الفرعية الضرورية للاختبار، وكذلك تقديم عدد كاف من الفقرات للمحافظة على صدق الاختبار.

اختبار بينية أول اختبار تكيفي :

يعد اختبار بينية أول اختبار تكيفي، لأنه يحتوي على معظـــم خـــصائص الاختبـــارات التكيفية الحديثة، فالفقرة الأولى التي تطبق على المفحوص تعتمد على قـــدرة المفحـــوص كمـــا يقدرها الفاحص قبل بدء الاختبار، والتصحيح يكون مباشر وتستخدم نتيجة هذا التـــصحيح فـــي تحديد الفقرات التي يليها، وتتوقف العملية عند الوصول إلى أعلى مستوى وهو المستوى الذي لا يستطيع المفحوص فيه الإجابة عن أية فقرة اختباريه بصورة صحيحة (Weiss, 1983) ويعـــد تيرن بول (Turn Ball) أول من استخدم لفظ تكيفي وقد كانت البداية في اختبارات ورقة وقلم، وقد اقترح (Lord, 1971) طريقة القياس التكيفي باستخدام نماذج النظرية الحديثة في القيـــاس، فهذه الاختبارات تختصر طول الاختبار إلى عدد فقرات اقل بينما الاختبـــارات التقليديـــة فـــإن فقراتها جميعها تقدم لجميع المفحوصين حيث يكون بعضها صعباً وبعضهاً سهلاً بينما يتم فـــي طريقة القياس التكيفي تشكيل اختبار خاص لكل مفحوص يتلاءم مع قدراته.

وقد أشار كرست (Crist, 1989) إلى الفوائد الآتية للاختبارات التكيفية:-

- الكفاءة والدقة العالية في تقدير قدرة المفحوص باستخدام عدد قليل من الفقرات المنسجمة مع قدرة المفحوص.
- المرونة العالية في الاختبارات المتمثلة في إمكانية تقديم نماذج مختلفة ومتنوعة من الاختبارات اعتمادا على قدرة المفحوصين.
- عدم وجود خطأ قياس ناجم عن ظروف تطبيق الاختبار خاصة إذا تم القياس التكيفي باستخدام الحاسوب.
- يعتمد تطبيق الاختبار على المفحوص نفسه، فليس هناك داع لأن ينتظر المفحوص زملاءه الآخرين لينتقل إلى إجابة الجزء التالي من الاختبار كما أن تعليمات الاختبار تقدم لكل مفحوص وهذا يوفر المزيد من الوقت للطلبة، ويقلل من قلق الاختبار.
- السرية العالية للاختبار.
- إمكانية قياس مدى واسع من المهارات أكثر مما توفره الاختبارات التقليدية (الاختبارات الكتابية).

أما هامبلتون وسوامينثان وروجرز (Hambleton, Swaminathan & Rogers, 1991) فقد أضافوا الفوائد الآتية:

الحصول على نتائج الاختبارات بشكل سريع وفوري عند استخدام الحاسوب.

أ. سهولة إزالة الفقرات التي تظهر فيها عيوب من بنك الأسئلة.

ب. المرونة في اختيار فقرات الاختبار.

ج. تقليل الإحباط الناجم عن تطبيق الاختبار.

وقد أشارت دانكل (Dunkel, 1999) إلى إمكانية استخدام وسائط متعددة في العــرض (Multimedia Presentation) تتضمن النصوص والصور والأفلام فــي القيــاس التكيفــي المحوسب، ويستخدم القياس التكيفي في تشخيص نقاط الضعف لبناء الخطط العلاجيــة، ونقــاط القوة لتعزيزها، إضافة إلى استخدامه في تحديد الحد الأدنى لمستوى الاتفاق في اختبارات التمكن (Mastery Tests) .

محددات القياس التكيفي باستخدام الحاسوب

بالرغم مما للقياس التكيفي باستخدام الحاسوب من فوائد وميزات عديــدة إلا أن هنــاك عددا من المحددات تحد من استخدامه، **وقد ذكر كرست (Crist, 1989) المحددات الآتية:**

- النقص في الأجهزة والمعدات والبرمجيات (Hardware And Software Limitation).
- تتطلب اختبارات القياس التكيفي دقة عالية في تدريج الفقرات (Item Calibration).
- تحتاج اختبارات القياس التكيفي باستخدام الحاسوب إلى ضرورة إتقــان الطلبــة لمهــارات استخدام الحاسوب.

كما أشار هامبلتون وسوامنثان (Hambleton & Swaminathan, 1985) إلــى الــصعوبات التالية التي تواجه الاختبارات التكيفية:

- صعوبة الإعداد والتحضير بسبب استخدام عدد كبير من الفقرات لتكوين تجمع الفقرات
- الكلفة المادية العالية من حيث الإعداد والتحضير وتجهيز البرمجيات .
- إمكانية التكرار لبعض الفقرات أحياناً.

أما كرست (Crist, 1989) فقد أشار إلى الكلفة المادية الخاصة بــالتجهيزات الماديــة (Hardware) وضرورة توخي الدقة العالية في عملية معايرة الفقــرات (Item Calibration) إضافة إلى ضرورة إتقان الطلبة لمهارات الحاسوب الأمر الذي لا يتوفر أحياناً.

أن الاختبارات التكيفية تعد ألان أداة متطورة في قياس السمات النفسية والتربوية بسبب دقة قياسها للسمة المقيسه ويُنتظر أن تكون اكثر شيوعا في المستقبل وبناءاً عليه فلا بــد مــن إجراء دارسات تفيد في تطوير الاختبارات التكيفية علــى جميــع المراحــل العمريــة للطلبــة وخصوصا المرحلة الأساسية الدنيا من السلم التعليمي في الوطن العربي والتي تشمل الــصفوف (الأول، الثاني، الثالث) كذلك يُنتظر تدريب المعلمين على الاســتراتيجيات المــستخدمة لتنفيــذ الاختبارات التكيفية لاستبدال الاستراتيجيات القديمة المملــة والمكلفــة فــي الوقــت والجهــد باستراتيجيات حديثة ومتطورة تواكب التطور العلمي وعصر التكنولوجيا لمساندة مشاريع وزارة التربية والتعليم والتي على رأسها مشروع اقتصاد المعرفة وحوسبة التعليم فلا بد مــن إجــراء دراسات من حيث استخدام الاختبارات التكيفية بأنواعها في نهاية كل فصل لقياس قدرات الطلبة بعد أن درسوا محتوى معد مسبقا لهم في وقت محدد وذلك لمعرفة مقدار ما تحقق من الأهداف لوضعهم في المكان المناسب لقدراتهم في شتى المجالات.

مشكلة الدراسة وأسئلتها

من الضروري توافر توازن معقول بين المرحلة المعرفية التي وصــل إليهــا الطالــب والمادة المقدمة له فالأسئلة السهلة جدا تؤدي إلى تعلم محدود، وأما الصعبة جــدا فتــؤدي إلــى الإحباط والفشل وبالتالي إلى إهمال الموضوع كاملا (Crist , 1989) وبظهور الاختبارات

التكيفية أصبح من الممكن انحسار هذه المشكلة والتغلب عليها باختبار الطالب بفقــرات مناسبة لقدرته حسب إحدى الاستراتيجيات المتبعة في الاختبارات التكيفية بدلاً من الاختبــارات التقليدية التي لا تراعي المدى الواسع من الفروق الفردية بين الأفراد على الرغم من أن المنهاج يفترض تقدم ثابت من المعرفة لدى طلبة الصف جميعهم ومن سنة إلى أخرى وعليه فالمنــاهج ليست متطابقة مع مستوى الفهم الذي يصل إلى العديد من الطلبة أو مرحلة نمائهم العقلي.

أما القياس المعاصر في القياس النفسي والتربوي فيسعى إلى أعلى مستوى من الدقة والموضوعية في القياس بحيث يتحقق اكبر قدر من الدقة في نتائج القياس وفي تقدير السمة الكامنة لدى الفرد وتحقيقا لهذا الهدف جاءت هذه الدراسة كمحاولة لبناء اختبار تكيفي لقياس القدرة الرياضية وفق الاستراتيجية ثنائية المرحلة في نظرية استجابة الفقرة لقياس القدرة الرياضية لطلبة الصف الثاني الأساسي في وحدتي الضرب والقسمة اللواتي يدرسن لأولئك الطلبة كخطوة على طريق بناء اختبارات تكيفية باستخدام الورقة والقلم أو باستخدام الحاسوب كبديل للاختبارات التقليدية التي تفترض تساوي قدرات الطلبة قبل البدء بتطبيق الاختبار وبالتحديد فإن هذه الدراسة سعت للإجابة عن التساؤلات التالية:

1- ما تقديرات الصعوبة لكل فقرة من فقرات الاختبار في عينة التدريج؟

2- ما تقديرات القدرة للأفراد (θ_1 , θ_2 ، θ_3) التي تم تقديرها باستخدام كل من الاختبار الاستطلاعي، الاختبار التكيفي الثنائي، جميع فقرات الاختبار؟

3- ما درجة اقتراب قيمة تقديرات القدرة في مرحلتي القياس التكيفي (θ_1 , θ_2) من تقدير القدرة عند تطبيق جميع فقرات الاختبار (θ_3)؟

4- ما مقدار معامل الارتباط بين قيمة(θ) في كل مرحلة من مراحل القياس ومحك العلامات المدرسية؟

مصطلحات الدراسة

نظرية الاستجابة للفقرة

نظرية حديثة في القياس النفسي والتربوي وأهم خصائصها أن تقدير قدرة المفحوص لا يعتمد على معالم الفقرات التي تقدم للمفحوص بالإضافة إلى تقدير معالم الفقرة يكون متحرراً من خصائص عينة المفحوصين الذين طُبق عليهم الاختبار. (Hambleton & Swaminathan, 1985)

الفقرة ثنائية التدريج

هي الفقرة التي يمنح الفرد عند استجابته استجابة صحيحة عليهـا العلامـة (1) أو أي علامة أخرى في حين يمنح العلامة صفر أو أي علامة أخرى عند اسـتجابته اسـتجابة خطـأ عليها.

القياس التكيفي

طريقة يتم فيها تطبيق الاختبارات بحيث يتم اختيار مجموعة من الفقرات لكل مفحوص تكون هي الأكثر ملاءمة لتقدير قدرته على أساس المطابقة بين قدرة المفحوص وصعوبة تلـك الفقرات، وفي هذه الدراسة تم استخدام الاستراتيجية ثنائية المرحلة كاستراتيجية من استراتيجيات القياس التكيفي.

الاختبار الاستطلاعي

اختبار يتكون من عدد قليل من الفقرات المدرجة يستخدم لتقـدير قيمـة أوليـة لقـدرة المفحوص.

الاستراتيجية ثنائية المرحلة

إحدى استراتيجيات القياس التكيفي التي تتضمن تقديم فقرات اختباريه مدرجة للمفحوص على مرحلتين:

- تقديم اختبار استطلاعي لتحديد مستوى القدرة الأولى.
- تعريض المفحوص لفقرات تناسب قدرته التي تحددت في الاختبار الاستطلاعي السابق.

صعوبة الفقرة (b_i)

أحد معلمات اقتران الاستجابة للفقرة وتعادل قيمة القدرة التـــي تنـــاظر احتمـــال 50% للإجابة الصحيحة عندما تكون قيمة التخمين تساوي صفر.

تمييز الفقرة (a_i)

هو ميل المماس لمنحنى خصائص الفقرة عند نقاط تقاطع منحنيـــات خـــصائص الفقرة.

تخمين الفقرة (c)

أحد معلمات اقتران الاستجابة للفقرة ويشير إلى احتمالية إجابة المفحـــوص ذي القـــدرة المتدنية على فقرة معينة إجابة صحيحة ،وهو القيمة على المحور الـــصادي المنـــاظرة لخـــط الاستواء السفلي (Lower Asymptote) في المنحنى اللوجستي للفقرة.

القدرة (θ):

مقدار ما يمتلكه الفرد من السمة موضوع القياس وقد توصل عبد العزيز (2006) إلى أن القدرة الرياضية بنيه مركبة من القدرات الفرعية (المجالات) التالية:

القدرة العددية، القدرة الاستدلالية، القدرة المكانية، القدرة على اكتساب المفاهيم،، وقد احتوى الاختبار الذي أعدته الباحثة على (27) فقرة تقيس القدرة العددية، كذلك (15) فقرة تقيس القدرة الاستدلالية، أيضا (7) فقرات تقيس القدرة المكانية و(15) فقرة تقيس القدرة على اكتساب المفاهيم وذلك حسب ما صنفه المحكمون للاختبار وبالتالي تألف الاختبار من (64) فقرة.

أن القدرة الرياضية هي القدرة على تفهم وامتلاك المفاهيم والرموز والمصطلحات والمهارات الأساسية وتوظيفها أو تطبيقها في مواقف رياضية جديدة كما أن القدرة الرياضية في تفكير اختياري في إطار العلاقات الكمية والتفسير الاستنتاجي والقدرة على تطبيق القواعد العامة على قاعدة الإعداد والرموز والأشكال الهندسية. ولأغراض هذه الدراسة فإن (θ) يعبر عن مقدرا امتلاك طالب الصف الثاني الأساسي في محافظة المفرق على القدرة الرياضية والقدرة الرياضية التي تناولتها هذه الدراسة هي القدرة على فهم وحل المسائل التي يعبر عنها بعمليات الضرب والقسمة وفقا لمستواه التعليمي المحدد مسبقا بالصف الثاني الأساسي.

أهمية الدراسة:

تأتي الأهمية العملية لهذه الدراسة من الانتشار الواسع لاستخدام القياس التكيفي في قياس الاتجاهات، والشخصية، والقدرات، والذي أصبح من الممكن بواسطته قياس قدرة الأفراد بطريقة دقيقة وبأقل كلفة من الجهد وبمعنى آخر بتعريض المفحوص لعدد من الفقرات اقل من عدد الفقرات في الاختبارات التقليدية وقد تطورت الاختبارات التكيفية من حيث الإعداد والتطبيق بفضل نظرية الاستجابة للفقرة وما قدمته من إمكانية لقياس معالم القدرة عند الأفراد بطريقة

متحررة من معالم الفقرات وكذلك إمكانية تقدير معالم الفقرات بطريقة متحررة من عينة التطبيق، هذا عدا عن المفاهيم الأخرى التي أهمها دالة المعلومات المرتبطة بكمية المعلومات التي تعطيها الفقرة عن المفحوصين، وكذلك ما يرتبط بها مما يسمى بالخطأ المعياري في التقدير والذي هو خاص بكل مفحوص بعكس مفهوم الخطأ المعياري في القياس في النظرية الكلاسيكية والذي يتعامل مع المفحوصين كمجموعة واحدة.

وتكمن أهمية مثل هذه الدراسة بتطبيق وسائل جديدة للقياس والتقويم التربوي عموما وللاختبارات التحصيلية على وجه التحديد أملاً في تطوير اختبارات يتم من خلالها وضع تقدير دقيق لقدرة الطالب، ولتشكل بالتالي أداة لتغذية راجعة يتم من خلالها تشخيص نقاط الضعف وتحديدها لبناء خطط علاجية مناسبة، وتحديد نقاط القوة لتعزيزها وجعلها مرتكزا للتعلم اللاحق . كذلك تزويد الباحثين ببيانات عن واقع الاختبارات التكيفية في المفرق والاستفادة من كيفية إجراء الدراسة والأداة المستخدمة فيها. كذلك تزويد الباحثين ببيانات عن واقع الاختبارات التكيفية في الأردن والاستفادة من كيفية إجراء الدراسة والأداة المستخدمة فيها.

كما تأتي الأهمية النظرية للدراسة من خلال إشارة أدب الموضوع إلى عدد من الاستراتيجيات في تطبيق القياس التكيفي الأمر الذي يعطي تقديرات مختلفة من حيث دقة القيمة المتحققة للقدرة ومن حيث درجة قربها أو بعدها عن القيمة الحقيقية للقدرة ومن هنا تأتي أهمية هذه الدراسة لبناء اختبار تكيفي لقياس القدرة الرياضية وفق الاستراتيجية ثنائية المرحلة وتطبيقه على طلبة المرحلة الأساسية الأولى في سلم التعليم الأكاديمي أملاً في أن تضيف للأدب التربوي شيئا جديدا يسهم في تحسين نواتج العملية التربوية.

يضاف إلى النقاط السابقة أنها تأتي كمحاولة للإجابة عن سؤال يتعلق بالاختبارات التكيفية وملاءمتها لطلبة المرحلة الأساسية الدنيا (الثاني الأساسي) كما أثبتت دراسات سابقة –

عرضت في الفصل الثاني- ملاءمتها للطلبة في مرحلة التعليم الأساسي المتوسطة والعليا وللصفوف الثانوية وطلبة الجامعات وغيرهم حيث إنه - وحسب حدود علم الباحثة- هذه الدراسة الأولى التي تعُنى بتكييف اختبار لطلبة المرحلة الأساسية الأولى من سلم التعليم الأكاديمي كذلك تطبيقها في منطقة واسعة تعد من المناطق الفقيرة والنائية في المملكة الأردنية الهاشمية منطقة محافظة المفرق بباديتها الشرقية والغربية بالإضافة للقصبة.

وتأتي هذه الدراسة منسجمة مع خطط وزارة التربية والتعليم في إيجاد تعلم مبني على اقتصاد المعرفة (Knowledge Based Economy)، الذي يسعى لتعلم متوافق مع متطلبات عصر تكنولوجيا المعلومات والاتصال (Information and Communication Technology)، و إدارة وتحليل البيانات والعمل بروح الفريق والانتقال بالتعلم من التلقين والسلبية إلى تعلم نشط مبني على الاستكشاف، والاستقصاء، ومن تعلم يعتمد على تلقي الحقائق وتخزينها إلى تعلم مبني على التفكير الناقد.

محددات الدراسة

* اقتصرت عينة القياس التكيفي على طلبة الصف الثاني الأساسي في المدارس المتعاونة والتابعة لمديرية التربية والتعليم في محافظة المفرق بباديتها الشمالية الغربية والشمالية الشرقية لعام 2007/2006 وهي مدرسة السعادة الغربية المختلطة الأساسية، مدرسة الثلاج الأساسية، مدرسة حي البوادي الأساسية، مدرسة روضة بسمة الأساسية، مدرسة البشرية الأساسية المختلطة.

* اقتصرت فقرات الاختبار على مجموعة فقرات أعدتها الباحثة وعليه فأن فاعلية هذه الفقرات تتحدد بدرجة مطابقتها للنموذج المستخدم في الدراسة ودلالات فاعليتها متمثلة بخصائصها الإحصائية.

*	اقتصرت فقرات الاختبارات التكيفية على قياس قدرات الطلبة في محتوى مادة وحـــدتي الضرب والقسمة للصف الثاني الأساسي للعام الدراسي 2007/2006م.

*	استخدم في الدراسة نموذج راش في تحليل بيانات الفقرات ويفترض هذا النموذج ثبات معلم التمييز لجميع الفقرات وان قيمة معلم التخمين تساوي صفر لجميع الفقرات.

*	استخدام برنامج (BIGSTEPS) الذي يعتمد طريقة الأرجيحة القصوى المشتركة فـــي تقدير معلم قدرة الأفراد وصعوبة الفقرات حيث أن هذه لطريقة تمكنا من تقدير قدرات الأفـــراد الذين أجابوا إجابة صحيحة أو إجابة خاطئة عن جميع فقرات الاختبار، وكذلك معالم الفقـــرات التي أجاب عنها جميع الأفراد إجابة صحيحة وإجابة خاطئة.

الفصل الثانـي

الدراسات السابقة

الفصل الثاني
الدراسات السابقة

يعتبر موضوع القياس التكيفي من المواضيع الحديثة التي لاقت قبولا كبيراً من قبل الباحثين فتعددت الدراسات حول هذا الموضوع ومنها من تناول القياس التكيفي ثنائي المرحلة ومنها من تناول القياس التكيفي متعدد المراحل بأنواعه وآخر دخل مجال القياس التكيفي المحوسب كذلك القياس التكيفي المبني باستخدام أسئلة ثنائية التدريج أو المبني باستخدام أسئلة متعددة التدريج وتم استعراض بعض من هذه الدراسات ذات الصلة بالموضوع. وفق التسلسل الزمني كما تم تصنيفها في فئتين هما:

الدراسات العربية

أجرى العجيلي (2000) دراسة بعنوان "استخدام الاستراتيجية ثنائية المرحلة في بناء اختبار تكيفي (Tailored) في قواعد اللغة العربية للصف السادس الابتدائي"، حيث استخدم الباحث اختبارا تكون من (90) فقرة في صورته النهائية لتقيس المستويات الثلاثة الأولى من تصنيف بلوم في قواعد اللغة العربية، وتألفت عينة الدراسة من (716) طالبا وطالبة اختيروا عشوائيا من مديريات التربية والتعليم الأربع في بغداد، وقد تراوحت صعوبة الفقرات بين (-3.15) و(3.28) لوجيت، بمتوسط مقداره (0.144) لوجيت وتراوح الخطأ المعياري بين (0.1409) و(0.3281) لوجيت، واستخدم الاختبار بعد تدريجه لبناء اختبار تكيفي وفق الاستراتيجية ثنائية المرحلة، ووزعت الفقرات على خمسة مستويات مختلفة في صعوبتها، يمثل كل منها مستوى من مستويات المرحلة الثنائية.

وقد تألف كل اختبار من اختبارات المرحلة الثنائية من (30) فقرة اختباريه تمثل المستوى المقيس بدقة عالية مع وجود تداخل بين فقرات المستويات المتتالية في الصعوبة كان حجمه

(15) فقرة، ولإعداد الاختبار الاستطلاعي (Routing Test) اختيرت فقرتان عشوائياً مــن كــل مستوى فرعي، فتكون الاختبار الاستطلاعي من (10) فقرات، حيث استخدم هذا الاختبار فــي التقدير الأولي لقدرة المفحوص وتوجيهه إلى اختبار المستوى الثاني المناسب لقدرته.

أما نور الدين (2002) فقد أجرى دراسة بعنوان "فاعلية اســتخدام الاختبــار المــوائم (التكيفي) باستخدام الحاسب في تقدير قدرة الأفراد وتحديد الخصائص السيكومترية للقياس" حيث تألفت عينة الدراسة لتقدير معالم فقرات الاختبار من (5000) طالب وطالبة من طلبــة الــسنة الثانية، والثالثة، والرابعة، وطلاب الدراسات العليا بكلية التربية بجامعة عــين شــمس، بينمــا تجاوزت عينة دراسة تطبيق الاختبارات التكيفية المحوسبة (5000) طالب وطالبة آخرين مــن تخصصات مختلفة، واستخدم الباحث (137) فقرة من اختبار رافــن للمــصفوفات المتدرجــة بصوره المختلفة في اشتقاق المعالم المختلفة للمفردات، وقام الباحث بتطبيق الاختبــار وتحليــل بياناته للتأكد من أحادية البعد، واختيار الفقرات الملائمة للنمــوذج، وأجــرى الباحــث تطبيقــا للاختبارات التكيفية المحوسبة من خلال ثلاث دراسات مختلفة، قارنــت الأولــى بــين فعاليــة الاختبار التكيفي المحوسب الذي يتوقف بعد تحقق درجة محددة من الدقة، مقارنــة بالاختبــار الخطي المحوسب، وقارنت الدراسة الثانية بين فعالية الاختبار التكيفي المحوسب الذي يتوقــف بعد عدد محدد من الفقرات مقارنة بالاختبار الخطي المحوسب أما الدراسة الثالثة فقد قارنت بين اختبارين تكيفيين محوسبين اســتخدم أحــدهما تقــدير القــدرة بطريقــة الارجحيــة العظمــى (Maximum Likelihood Estimation) واســـتخدم الآخـــر الطريقـــة البايزيـــة (Bayesian Modal Estimation) وكل منهما ينتهي بتطبيق عــدد محــدد مــن الفقــرات، وأشارت نتائج الدراسة إلى أن الاختبار التكيفي يعطي دقة أعلى من الاختبار الخطي لأنه ينتقي الفقرات المناسبة لقدرة الفرد، كما انه اقصر حيث يمكن الوصول إلى دقة أعلى من خلال تطبيق

(25) فقرة، بينما احتاج الاختبار الخطي المحوسب إلى (60) فقرة وتجاوز معامل ثباته (0.80) من خلال تطبيق (5) فقرات فقط، كما وصل معامل ثبات الاختبار التكيفي المحوسب إلى معامل ثبات مقداره (0.94) من خلال تطبيق (50%) من الفقرات، كما أشارت النتائج إلى أن أسلوب تقدير القدرة باستخدام الطريقة البابيزية اكثر دقة من الاختبار الذي استخدم أسلوب الارجحية العظمى في تقدير القدرة حيث انه يقدم معلومات اكثر، وله خطأ معياري اقل من الارجحية العظمى، إلا أن أسلوب الأرجحية العظمى أعطى متوسطا أعلى لتقدير القدرة مقارنة بالطريقة البابيزية. وتوصل الباحث إلى أن الاختبار التكيفي المحوسب امتاز بقلة عدة فقرات الاختبار بنسبة تزيد عن (50%) في المتوسط مقارنة بالاختبار الخطي، وزاد الاختبار التكيفي من دقة القياس بنسبة وصل متوسطها إلى (20%) مقارنة بالاختبار الخطي، إضافة إلى أن الاختبار التكيفي أعطى تقديرا للقدرة لا يختلف متوسطه عن تقدير القدرة للاختبار الخطي.

كذلك الدراسة التي أجراها العموش (2003) بعنوان "فاعلية القياس التكيفي في تقويم بعض القدرات المعرفية لدى طلبة السنة الأولى الجامعية" وهدفت هذه الدراسة لاستقصاء فاعلية القياس التكيفي في تقويم القدرة اللفظية والقدرة الرياضية لدى طلبة السنة الأولى الجامعية ولذلك قام الباحث بتطوير اختبارين في القدرة اللفظية والقدرة الرياضية وفق الطريقتين التقليدية والحديثة، ومن ثم دراسة الخصائص السيوكومترية للقياس التكيفي وتطبيق الإختبارين على عينة التدريج وذلك لمعايرة الفقرات ومطابقة النموذج الثلاثي المعلمة المستخدم في الدراسة. وأظهرت نتائج الدراسة تفوق القياس التكيفي لاختبار القدرة الرياضية في مستويات القدرة المنخفضة ومستوى القدرة المرتفعة، أما في الاختبار التكيفي للقدرة اللفظية فقد كان الاختبار التقليدي اكثر فاعلية من القياس التكيفي في مستويات القدرة المتوسطة والمرتفعة. كما أظهرت نتائج الدراسة أن القياس التكيفي كان أكثر فاعلية من الاختبارات التقليدية وذلك اعتماداً على مؤشرات دقة

القياس مقدرة بالخطأ المعياري في تقدير القدرة، والكفاءة النسبية واختزال الوقت، وعدد الفقرات المستخدمة.

وهدفت دراسة الثوابية (2004) إلى استقصاء فاعلية القياس التكيفي ثنائي المرحلة في تقويم التحصيل في مبحث الأحياء لدى عينة من طلبة الصف الثاني الثانوي العلمي في الأردن حيث قام ببناء اختبار تحصيلي في وحدة الوراثة من كتاب الأحياء للصف الثاني الثانوي العلمي تكون من (55) فقرة ثنائية التدريج من نوع الاختيار من متعدد وطبق هذا الاختبار على عينة مكونة من (721) طالباً وطالبة ودرجت فقراته باستخدام البرنامج الإحصائي (Winsteps) اعتماداً على معلم الصعوبة وتم اشتقاق ستة اختبارات فرعية أحدها اختبار المرحلة الأولى الذي يطلق عليه الاختبار الاستطلاعي والذي يتم من خلاله توجيه المفحوص إلى واحد من الاختبارات الخمسة الأخرى التي تكون المرحلة الثانية من الاختبار التكيفي وتم حساب ثلاثة تقديرات لقدرة كل مفحوص أولى هذه التقديرات تم الحصول عليها باستخدام فقرات الاختبار الاستطلاعي وثاني هذه التقديرات تم الحصول عليها من الاختبار الذي شكل المرحلة الثانية مضافا إليها فقرات الاختبار الاستطلاعي والتقدير الثالث للقدرة تم الحصول عليه من جميع فقرات الاختبار الخمس والخمسين وقد رافق ذلك في كل مرة حساب الخطأ المعياري في التقدير لكل مفحوص.

وجرى بعد ذلك مقارنة للقدرة المحسوبة باستخدام الاختبار الاستطلاعي مضافاً إليه اختبار المرحلة الثانية وكذلك القدرة المقدرة من جميع فقرات الاختبار (باستخدام الاختبار الخطي غير التكيفي) حيث كانت تقديرات الاختبار التكيفي للقدرة بشكل عام أعلى من تقدير الاختبار غير التكيفي إلا أن الخطأ المعياري في تقدير القدرة كان اقل باستخدام جميع الفقرات وقد بلغ معامل الارتباط بين القدرة المحسوبة من الاختبار الاستطلاعي والعلامة المدرسية

(0.75) وبينها وبين علامة الثانوية العامة (0.46) وبلغ معامل الارتباط بين القدرة المحسوبة من الاختبار التكيفي ككل والعلامة المدرسية (0.70) وبينها وبين علامة الثانوية العامة (0.42) أما معامل الارتباط بين القدرة المقدرة من جميع فقرات الاختبار والعلامة المدرسية (0.76) وبينها وبين علامة الثانوية العامة (0.52).

وفي دراسة ماري فيسر (Vesser, 1998) في جامعة جنوب إفريقيا والتي قارنت فيها الطريقة التقليدية (الورقة والقلم) وطريقة القياس التكيف وطريقة اختبار بواسطة الحاسوب لاختيارات القدرات المعرفية: (اللفظية وغير اللفظية) لطلبة المرحلة الثانوية في اختيار الاستعداد (G-SAT) من حيث دقة القياس وعدد الفقرات المستخدمة والوقت اللازم حيث استخدمت النموذج الثلاثي المعلم في النظرية الحديثة (IRT) لمعايرة وتدريج الفقرات واستخراج معالمها وكما استخدام برنامج محوسب في تطبيق الاختبارات بواسطة الحاسوب وكانت فقرات الاختبار من نوع الاختيار من متعدد، وتم اختيار (15) فقرة لكل من الاختبار اللفظي وغير اللفظي والاختبارات المبرمجة، وأشارت نتائج الدراسة إلى أن نتائج الطلبة كانت أفضل بواسطة الطريقة التقليدية منها في القياس التكيفي أو الاختبارات المبرمجة مع أن الخصائص السيكومترية للثبات والصدق متماثلة في الاختبارات الثلاثة، وقد يعود ذلك لتأثر استخدام الحاسوب على الطلبة، ووجد أن القياس التكيفي يوفر 75% من الوقت مقارنة مع الاختبارين الآخرين وكذلك تم استخدام (15) فقرة فقط في القياس التكيفي.

ومن الدراسات الحديثة في موضوع القياس التكييفي الدراسة التي أجراها البرصان (2006) حيث هدفت هذه الدراسة التي استقصاء اثر عدد مراحل القياس التكييفي وعدد أسئلة كل مرحلة في تقدير القدرة والخطأ المعياري في التقدير ولتحقيق أهداف الدراسة تم بناء اختبارين تحصليين لوحدتي الإعداد الحقيقية والأسس واللوغاريتمات تكون الأول من (50) فقرة

ذات استجابة ثنائية التدريج، وتكون الثاني من (21) فقرة ذات استجابة متعددة التـدريج حيـث طبق الاختبار ذو الفقرات ثنائية التدريج على عينة مكونة من (408) طالباً وطالبة واسـتخدمت بيانات إجابات الطلبة في تدريج فقرات كل من الاختبارين حسب نمـوذج راش فـي نظريـة الاستجابة للفقرة باستخدام البرنامج الإحصائي (Winsteps) وقد قسمت فقرات الاختبـار الأول ثنائية التدريج إلى اختبار استطلاعي شكل المرحلة الأولى من القياس التكيفي وخمسة اختبارات فرعية شكلت المراحل من الثانية إلى الرابعة وطبقت تتابعياً بحيث يتحدد الاختبار الفرعي فـي أي مرحلة بناء على القدرة المتحققة من الاختبار الفرعي الذي شكل المرحلة السابقة.

أما فقرات الاختبار الثاني متعددة التدريج فقد قسمت كذلك إلى اختبار استطلاعي شـكل المرحلة الأولى من القياس التكيفي وأربعة اختبارات فرعية شكلت المراحل الثانية حتى الرابعة، وقد تم الحصول للمفحوص سواء في حالة الاختبار ذي الفقرات ثنائية التـدريج أو فـي حالـة الاختبار ذي الفقرات متعددة التدريج على خمسة تقديرات للقدرة هي (θ_1) باستخدام الاختبـار الاستطلاعي، (θ_2) باستخدام الاختبار الفرعي الذي شكل المرحلة الثانية، (θ_3) باسـتخدام اختبار المرحلة الثالثة، (θ_4) باستخدام اختبار المرحلة الرابعة (θ_5) باسـتخدام جميـع فقـرات الاختبار الخمسين ثنائية التدريج أو باستخدام جميع فقرات الاختبار الإحدى والعشرين متعـددة التدريج.

وقد رافق كل تقدير للقدرة إيجاد الخطأ المعياري في تقدير قدرة المفحوص، وللكـشف عن اثر عدد مراحل القياس التكيفي ثم استخدام معيارين هما معامل الارتباط ومتوسطات الفروق بين القدرة المقدرة في كل مرحلة والقدرة المرجعية المقدرة في المرحلة الخامسة باستخدام جميع فقرات الاختبار وبالنسبة للخطأ المعياري فقد تم استخدام معيار متوسط الفـروق بـين الخطـأ

المعياري في تقدير القدرة في كل مرحلة والخطأ المعياري في المرحلة الخامسة (المرجعية) وتوصلت الدراسة إلى النتائج التالية:

- يستمر تقدير القدرة والخطأ المعياري في التقدير في كل مرحلة من تقدير القدرة المرجعية والخطأ المعياري المرجعي عند زيادة عدد المراحل إلا أن الزيادة لم تعد مؤثرة بعد المرحلة الثالثة في حالة القياس التكيفي المعتمد على فقرات ثنائية التدريج ولم تعد مؤثرة بعد المرحلة الثانية في حالة القياس التكيفي المعتمد على فقرات متعددة التدريج.
- كان عدد فقرات المراحل الأفضل حسب معامل الارتباط ومتوسطات الفروق بالنسبة لتقدير القدرة هو (6) فقرات في حالة القياس التكيفي المعتمد على فقرات ثنائية التدريج و(3) فقرات في حالة القياس التكيفي المعتمد على فقرات متعددة التدريج.
- كان القياس التكيفي المبني على فقرات متعددة التدريج أدق من حيث تقدير القدرة والخطأ المعياري في التقدير من القياس التكيفي المعتمد على فقرات ثنائية التدريج.

الدراسات الأجنبية

وفي دراسة لورد (Lord, 1971a) التي اقترح فيها استراتيجية لتحويل اختبار تقليدي إلى اختبار تكيفي أطلق عليه اسم اختبار مرن المستوى (Flexi Level Test) باستخدام الورقة والقلم في محاولة منه لتطبيق نظرية السمات الكامنة وذلك بربط مستوى صعوبة الفقرة بمستوى قدرة المفحوص. وفي هذه الاستراتيجية يطلب من المفحوص أولا الاستجابة إلى فقرة متوسطة الصعوبة، فان كانت إجابته صحيحة تطرح عليه فقرة اصعب، وان كانت خاطئة تقدم له فقرة اسهل وهكذا حتى ينتهي الاختبار. ونشر لورد الدراسة التي حول فيها اختباراً تحصيلياً تقليدياً إلى اختبار مرن المستوى بعد أن طبقه على عينة تألفت من (119) فرداً. وتوصل من خلال هذا

التطبيق إلى أن الاختبار مرن المستوى يعطي نتائج قياس أكثر دقة عنـــد المـــستويين العـــالي والمتدني من القدرة المقيسة فضلاً عن اختصار عدد الفقرات التي يجيب عنها المفحـــوص فـــي الاختبار .

أما دراسة لأركن وويز (Larkin & Weiss, 1975) فقد هدفت هذه الدراســـة إلـــى مقارنة اختبار تكيفي مبني وفق الاستراتيجية الهرمية متعددة المراحل واختبـــار تكيفـــي وفـــق الاستراتيجية ثنائية المرحلة. ولأغراض الدراسة طبق الباحثان اختبارا هرميا تألف مـــن (15) مرحلة بواقع عدد من الفقرات مساوي لرقم المراحل، وذلك لمراحل الاختبارية جميعها، ويجيب المفحوص عن فقرة اختبارية واحدة في كل مرحلة واختبار ثنائي المرحلة تألف من (40) فقرة لكل اختبار من اختبارات المرحلة الثانية، وقد تألفت عينة الدراسة من (111) طالبا جامعيا فـــي المرحلة الجامعية الأولى في مينسوتا. وأظهرت النتائج أن توزيع الدرجات في كلا الاختبـــارين كانا ذات التواء موجب؛ ولكنه لا يختلف اختلافا دالا إحصائيا عن التوزيع الاعتـــدالي، وكـــان توزيع الدرجات في الاختبار الهرمي مفرطحا (platykurtio) في حين كان التوزيـــع محـــدباً (leptokurtic) في الاختبار ثنائي المرحلة . كما اتضح أيضا أن تعيين الفـــرد علـــى اختبـــار المرحلة الثانية يكون أكثر دقة مما هو في المرحلة التوجيهية في الاختبار ثنائي المرحلة، إذ أن خطأ التحديد كان بنسبة اقل من (1%) وكان الارتباط بين درجات الاختبار الهرمي والاختبـــار ثنائي المرحلة على مدى تراوح بين (0.79) و(0.84) فكلا الاستراتيجيتين يبدو أنهـــا تكيـــف صعوبة الفقرات لتناسب مستوى المفحوص ؛مما يقلل من اثر التخمين في العمليـــة الاختباريـــة ولكن يبدو أن الإستراتيجية الهرمية أكثر نجاحا بقليل من الإستراتيجية ثنائية المرحلة في خفض عامل التخمين.

وأجرى هالادينا ورويد (Haladyna & Royd , 1983) دراسة بعنوان "مقارنة بين طريقتين لبناء الاختبارات التكيفية محكية المرجع "هدفت إلى المقارنة بين طريقتين لاختبار فقرات الاختبارات التكيفية هما الاختيار العشوائي للفقرات، والاختيار المبني على صعوبة الفقرات، وقد طبق الباحثان أربع استراتيجيات لاختيار الفقرات الاختبارية، التي قدمت لطلبة مادة تحليل الأسنان بكلية طب الأسنان وتم بناء ثلاث اختبارات تكيفية مطابقة لنموذج راش وبمستويات صعوبة مختلفةً.

وقد كانت الاختبارات مختلفة الطول، إذ كان عدد فقراتها (10) و (20) و(30) و(40) فقرة على التوالي. وبشكل عام كان الخطأ المعياري للأختبارات المحكية التكيفية في الطريقتين الأولى والثانية أقل من الخطأ المعياري المتحقق وفق الطريق الرابعة وهي الاختيار العشوائيٍ، وقد كان الاختبار الأفضل هو الذي يتضمن فقرات تناسب قدرة المفحوص حيث كان الخطأ المعياري أقل من الطرق الأخرى، وقد كان الخطأ المعياري في الطريقة الرابعة وهي الاختبار العشوائي أقل من الطريقة الثالثة وهي إعطاء فقرات صعوبتها بعيدة عن قدرات المفحوص.

وفي دراسة لبرنجستروم ولوز وجنشون (Brengstrom, Luhz & Genshon, 1992) بعنوان " تغير مستويات الصعوبة في الاختبار التكيفي المحوسب" هدفت لاختبار أثر تغير مستويات الصعوبة على قياس القدرة وطول الاختبار التكيفي المحوسب حيث تم تطبيق الدراسة على (225) مفحوصا تعرضوا عشوائياً لثلاث حالات من الصعوبة (مرتفعة، متوسطة، متدينة) حيث ثم اعتماد الاحتمالات للإجابة الصحيحة (50%، 60%، 70%) أظهرت النتائج بأنه لا تأثير لتغير احتمال الإجابة الصحيحة على تقدير قدرة المفحوص، إضافة إلى إن الاختبار ألتكيفي المحوسب السهل جعل عدد الفقرات الضرورية يزيد بشكل طفيف حتى يصل قاعدة التوقف.

وأجرى كيم وبليك (Kim & Plake, 1993) دراسة بعنوان " مقارنة بين اختبار تكيفي ثنائي المرحلة واختبار تكيفي محوسب باستخدام طريقة المحاكاة " هدفت إلى إجراء مقارنة بين اختبار تكيفي محوسب واختبار تكيفي ثنائي المرحلة، ولتحقيق الهدف تم توليد استجابات (1600) مفحوص باستخدام المحاكاة على ثلاثة اختبارات تكيفية محوسبة مختلفة في عدد الفقرات فقط، واختبار ثنائي المرحلة تألفت مرحلته الثانية من سبعة اختبارات فرعية كل منها تحتوي (18) فقرة، أظهرت النتائج أن الخصائص الإحصائية (صعوبة الفقرات) للاختبار الاستطلاعي لها اثر أساسي في دقة قياس القدرة وبشكل عام تفوق الاختبار التكيفي المحوسب على الاختبار ثنائي المرحلة المساوي له بعدد الفقرات وفق محكي دقة القياس والفعالية، كذلك وجد أن الاختبارات ثنائية المرحلة التي استخدمت نموذج راش بمدى واسع من الصعوبة في المرحلة الأولى، تعطي تقديرات اكثر دقة من الأشكال الأخرى للاختبارات ثنائية المرحلة و أشار الباحثان إلى أن الاختبارات ثنائية المرحلة تعتبر بديلا عمليا مناسبا عند عدم توافر أجهزة الحاسوب.

وفي دراسة لشنبكي وجرين (Schnipke & Green, 1995) بعنوان "مقارنة بين طرق اختيار فقرات الاختبارات الخطية والاختبارات التكيفية " هدفت إلى مقارنة طرق اختيار فقرات القدرة المعرفية باستخدام الاختبارات التكيفية والاختبارات الخطية، حيث اعتمد اختيار الفقرات في الاختبار الخطي على معيار القدرة التمييزية العالية، أما اختيار فقرات الاختبار التكيفي فقد تم اعتماد معيار دالة المعلومات بحيث يتم اختيار الفقرة التي تعطي اكبر كمية من المعلومات، واستخدم الباحثان عينة تكونت من (1000) مفحوص، ثم اختبارهم بمجموعة فقرات تقيس القدرة المعرفية حيث تكونت من (4) أو (5) أو (7) فقرات من اصل (15) و(30) و (50) فقرة على التوالي، وأظهرت النتائج أن الاختبارات التكيفية أعطت اكبر قدر من المعلومات عن قدرة

المفحوص مقارنة بالاختبارات الخطية، كما أن مربع الخطــأ المعيــاري لتقــدير القــدرة فــي الاختبارات التكيفية اقل منه في الاختبارات الخطية.

وأجرى شيرمز وفلكرسون (Shermis & Fulkerson, 1996) دراسة كــان هــدفها استخدامات اختبارات تكيفية محوسبة في الرياضيات حيث تم بناء اختبار تكيفي محوسب لوضع طلبة الصف الخامس الابتدائي في مدرسة خاصة للموهوبين في الرياضيات، ولهذه الغاية تم بناء بنك أسئلة تكون من (240) فقرة مثلت(60) هدفاً، وقسمت هذه الفقرات إلــى ثمانيــة نمــاذج اختباريه، واستخدمت ثلاث عينات لفحص فاعلية الاختبار التكيفي المحوسب في الرياضــيات: العينة الأولى تكونت من (683) طالبا من طلبة الصف السادس اختيروا من مدارس المنــاطق الريفية وضواحي المدن، وسبب اختيار طلبة الصف السادس هو وجود فقرات تعتمد على منهاج هذا الصف، وتكونت العينة الثانية من (190) طالبا من طلبة الــصف الــسادس فــي مــدراس الموهوبين في الرياضيات، وتكونت العينة الثالثة من(199) طالبا من طلبة الــصف الخــامس اختيروا من أربع مدارس مختلفة من نفس المنطقة التعليمية وهم من الطلبــة الــذين رشــحهم معلموهم ليكونوا ضمن مدارس الطلبة المتفوقين في الرياضيات.

ولتحقيق أهداف هذه الدراسة تم تقسيم عينة الدراسة إلى مجموعتين حيث تم تقديم ثمانية نماذج من اختبارات كتابية للمجموعة الأولى في حين استخدم الاختبار التكيفي المحوسب لقياس قدرات المجموعة الثانية، وأظهرت نتائج الدراسة أن متوسط معامل ثبات الاختبارات الكتابية بلغ (0.76) ومتوسط تقدير قدرة المفحوصين (0.45) لوجيت بانحراف معياري مقــداره (0.66)، وبلغ معامل الارتباط بين الاختبار التكيفي المحوسب وتقديرات المعلمين (0.19) كما أظهــرت النتائج أن الاختبار التكيفي المحوسب اكثر فاعلية في اختيار الموهوبين من تقديرات الاختبارات الكتابية.

ومن الدراسات التي اهتمت بتحويل الاختبار التقليدي إلى اختبار تكيفي محوسب دراسة شيرمز وستمر (Shermis & Stemmer, 1996) والتي تم فيها تحويل اختبار للرياضيات للصف العاشر في ولاية ميتشغان من الطريقة التقليدية (الاختبار الكتابي) إلى اختبار تكيفي وتم استخدام نموج راش أحادي المعلمة في تدريج الفقرات. وتألف هذا الاختبار من خمسة اختبارات فرعية: كل منها يتكون من خمس وعشرين فقرة، وطبقت هذه الاختبارات على طلاب الصف العاشر في مدارس مختلفة حيث تألفت العينة من (500) طالب من مستوى الصف التاسع إلى الصف الثاني عشر وتم استخدام (19) فقرة لأغراض القياس التكيفي وأظهرت نتائج الاختبارات أن معامل الثبات بطريقة الاتساق الداخلي كان مقبولا ما عدا الاختبار الرابع حيث بلغ معامل ثباته (0.58) وتم التحقق من افتراضات نظرية استجابة الفقرة (أحادية البعد، والاستقلال الموضوعي، منحنيات خصائص الفقرة). كما تم حساب صعوبات الفقرة حيث تراوحت بين (−0.32ـ1.44) . وكان طول الاختبار التكيفي يقل بنسبة (25%) عن الاختبار الكتابي، كما استخدمت طريقتا الأرجحية العظمى والطريقة البابيزية لتقدير المعالم ووجد أن الطريقة البابيزية مناسبة بشكل يماثل الأرجحية العظمى؛ ولكنها تحتاج لتحديد مسبق لقدرة العينة القبلية، ووجد أن (75%) من الطلبة اظهروا ارتياحا للقياس التكيفي، كما تم توفير (70%) من الوقت مقارنة مع الاختبار التقليدي .

أما دراسة كاسل (Castle, 1997) فقد هدفت لاستقصاء الكفاءة النسبية للقياس التكيفي ذي المرحلتين مقارنة مع اختبار تقليدي، وتم معايرة فقرات الاختبار باستخدام النموذج ثلاثي المعالم حيث اختيرت (180) فقرة بعد تدريجها واستخرج معالمها وتم استخدام (13) فقرة في اختبار المرحلة الأولى (Routing) واختيار (30) فقرة في المرحلة الثانية وكانت المتغيرات في الدراسة طول الاختبار في المرحلة الأولى والثانية وعدد الفقرات المستخدمة وخطأ القياس

ومؤشر الثبات وتوزيع مستويات القدرة ودالة المعلومات وأظهرت النتائج أن الثبات ودقة القياس والكفاءة النسبية كانت عالية في القياس التكيفي وكذلك قصر الاختبار مقارنـــة مـــع الاختبـــار التقليدي بنسبة (20%).

وأجرى الخضر وكلارك(Alkadher and Clarke , 1998) دراسة للبحث في تكافؤ اختبارات الاستعداد التفاضلي التكيفي المحوســب (CAT of the Differential Aptitude Test CAT DAT) واختبارات الورقة والقلم، ولهذه الغاية تم استخدام ثلاثة اختبارات؛ الأول اختبار القدرة العددية والثاني اختبار التفكير المجرد والثالث اختبار التفكيـــر الميكـــانيكي، وقـــد أجريت الدراسة على (122) من العمال المتدربين في شركة تكرير نفط تراوحت أعمارهم بين (18–20)، حيث وزعوا عشوائيا إلى مجموعتين: المجموعة الأولى تـــم اختبـــارهم باســـتخدام اختبارات كتابية والمجموعة الثانية تم اختبارهم باستخدام اختبارات تكيفية محوسبة وقد أظهرت نتائج الدراسة تكافؤ كلتا الطريقتين في قياس القدرة العددية، والتفكير الميكانكي بينما كان الأداء اكثر ارتفاعا على الاختبار التكيفي المحوسب للتفكير المجرد من الاختبار التقليدي، وقد وجد بأن العلاقة بين كل من التفكير المجرد والتفكير الميكانيكي والتحصيل بشكل عام علاقة إيجابية ذات دلالة إحصائية حيث بلغت قيمة معامل الارتباط ما بين 0.27 –0.41، وكذلك وجد بأن الارتباط بين التحصيل العام لأفراد العينة واختبارات القياس التكيفي أعلى من الارتباط بـــين اختبـــارات الورقة والقلم والتحصيل العام.

أما دراسة بيرجوف (Burghof ,2001) فقد هدفت لبناء بنك فقرات من نوع اختيـــار من متعدد في مبحث الجغرافيا لاستخدام فقرات البنك فـــي بنـــاء اختبـــارين أحـــدهما خطـــي محوســـــب(Linear On-Line Test, LOT) والآخـــــر تكيفـــــي محوســـــب (Computerized Adaptive Test, CAT)، فقامت الباحثة بإعداد جدول مواصفات (Blue

(Print لمبحث الجغرافيا للصف الثامن الأساسي مستخدمة خبرات المعلمين في هـذا المجـال، حيث قسم المحتوى إلى (13) موضوعاً فرعياً، وصنفت الأهداف التعليمية حسب تصنيف بلـوم للأهداف المعرفية. بعد ذلك تم تقدير عدد فقرات البنك بحيث لا تقل الفقرات عن 5 فقـرات ولا تزيد عن (10) فقرات لكل مفحوص، وذلك لتفادي تكرار الفقرات وكشفها لاسيما في الاختبارات الخطية، ثم قدرت الباحثة توزيع الفقرات على المواضيع الفرعية المختلفة، بعد ذلك تمت كتابـة (375) فقرة من نوع اختيار من متعدد ووزعت الفقرات على نموذجين (أ، ب) بواقع (54) فقرة لكل نموذج بحيث يشترك النموذجان بـ 10 فقرات كجذع مشترك، بعد ذلك طبـق النموذجـان على عينتين مختلفتين من المفحوصين استخدمت البرمجية الحاسوبية (ConQuest) في معادلة معالم الفقرات بين النموذجين باستخدام نموذج راش الأحادي المعلم، بعد ذلك خزنت الفقرات في قاعدة بيانات باستخدام البرمجية الحاسوبية (Microsoft Access 2000) بحيث تخزن الفقرات مع بيانات مثل: رقم الفقرة، موضوعها، نوعها، السنة الدراسية، الفصل الدراسي، صعوبتها .... ثم استخدمت فقرات البنك في سحب اختبارين محوسبين أحدهما خطي والآخر تكيفي في مبحث الجغرافيا للصف الثامن.

وقام هندركسون (Hendrickson, 2002) بدراسة بعنوان "تدريج اختبـار تحـصيلي ثنائي المرحلة" وكان هدف هذه الدراسة الكشف عن إمكانية استخدام الاختبار التكيفـي ثنـائي المرحلة بديلا عن الاختبارات التحصيلية التقليدية وتكونت أدوات الدراسة من اختبار في اللغـة الإنجليزية اشتمل على المفردات واستخدام اللغة والاستيعاب القرائي واختبار رياضيات اشـتمل على حل المشكلات وتفسير البيانات واختبـار (IOWA) للمهـارات الأساسـية واسـتخدمت الاستراتيجية ثنائية المرحلة في بناء الاختبار التكيفي وطبق الاختبار على طلبة الصف السادس واستخدم الخطأ المعياري للقياس وعلامات القطع ودالة المعلومـات كمحكـات للمقارنـة بـين

الاختبار التقليدي والاختبار التكيفي ثنائي المرحلة وبينت نتائج الدراسة أن الخصائص السيكومترية المتعلقة بالثبات ودقة القياس تتحسن عند تدريج الاختبارات لتصبح اختبارات تكيفية.

وفي دراسة لـ ديربير وفيزر (De Beer&Visser,1998) تم مقارنة اختبار الاستعداد المدرسة العام (GSAT) التكيفي المحوسب واختبار الورقة والقلم حيث تكونت عينة الدراسة من (20) مدرسة، واختبر من كل مدرسة (18) طالباً و(18) طالبة وكان الاختبار عشوائيا للمدارس والطلبة وكانت أعمار الطلبة (18)عاما، وتكونت أداة القياس من اختبارين متكافئين لـ (GSAT) من نوع اختبارات الورقة والقلم حيث استخدم لبناء اختبار تكيفي محو سب. وهدفت هذه الدراسة إلى فحص تكافؤ نتائج هذين النوعين من الاختبارات، حيث قدم لكل مفحوص إحدى صورتي اختبار الورقة والقلم والاختبار التكيفي المحوسب أولا من ثم اختبار الورقة والقلم والعكس بالنسبة للمجموعة الثانية.

وأشارت نتائج الدراسة إلى تكافؤ تحصيل الطلبة على نموذجي الاختبار حيث كان تحصيلهم على (اختبار الورقة والقلم أعلى من تحصيلهم) على (الاختبار التكيفي المحوسب)، وقد عزى الباحث هذا التباين في النتائج إلى أن الطلبة أكثر ألفة باختبارات الورقة والقلم مقارنة بالاختبارات التكيفية المحوسبة.

يمكن تصنيف الدراسات السابقة حسب أهدافها إلى سته أنواع فمنها :

1. دراسات كان هدفها بناء اختبارات تكيفيه مثل دراسة العجيلي (2000).
2. دراسات تهدف إلى تحويل اختبارات تقليدية إلى اختبارات تكيفيه مثل دراسات لورد (Lord, 1971a) وشيرمز وستمر (Shermis & stemmer, 1996) وهندركسون (Hendrickson, 2002).

3. دراسات تهدف للمقارنة بين استراتيجيات الاختبارات التكيفية مثل دراسة لاركن وويــز (Larkin & Weiss, 1975). وهالا دينا ورويد (Haladyna & Royd, 1983) وكــيم وبليك (Kim & Plake, 1993) وشنبكي وجرين (Schinpke & Green, 1995) وكاسل (Castle, 1997) والخضر وكلارك (Alkadher and Clarke, 1998).

4. دراسات تهدف إلى دراسة فاعلية الاختبارات التكيفية مثل دراسات نور الدين والعموش (2003) والثوابية (2004).

5. دراسات استخدمت الاختبارات التكيفية وظيفيا مثــل دراســات شــيرمز وفلكرســون (Shermise & Fulkerson, 1996).

6. دراسات تهدف لبناء بنك فقرات لمواد مختلفة واستخدام تلك الفقرات لبناء اختبار تكيفي مثل دراسات الفرجات (2004) وبيرجوف (Burghof, 2001).

ملخص الدراسات السابقة

من الدراسات السابقة يمكن إجمال بعض النتائج العامة منها :

- معامل ثبات الاختبار التكيفي المحسوب أعلى من معامل ثبات الاختبار التقليدي كما في دراسة شيرمز وستمر (Shermis & stemmer, 1996) ودراسة شيرمز وفلكرســون (Shermise & Fulkerson, 1996) ودراسة شنبكي وجرين (Schinpke & Green, 1995).

- إن القياس التكيفي يعطي تقديرا أعلى لقدرة المفحوص مقارنة بالاختبار الكتــابي وذلــك بالاعتماد على نتائج دراسة الخضر وكلارك (Alkadher and Clarke, 1998).

- إن القياس التكيفي أكثر دقة في تقدير قدرة المفحوص من الاختبارات الكتابية التقليديـــة كما في دراسة هندركسون (Hendrickson, 2002) وشيرمز وفلكرسون (Shermise & Fulkerson, 1996) ونور الدين (2002).
- إن القياس التكيفي يعمل على تقليل عدد فقرات الاختبار اللازمة لتقدير قدرة المفحـــوص وبالتالي يعمل على توفير الوقت كما في دراسة نور الدين (2002).
- بالنسبة للمطابقة بين الاختبارات التكيفية ثنائية المرحلة والاختبارات التكيفيـــة الهرميـــة فتشير دراسة لاركن وويز (Larkin & Weiss, 1975) إلى تفوق الاختبارات الهرمية من ناحية الدقة في تقدير القدرة.

ولكن خلت تلك الدراسات من دراسات في الاختبارات التكيفية لطلبة المرحلة الأساســية الأولى من سلم التعليم الأكاديمي طلبة الصفوف (الأول، الثاني، الثالث) الأساسي كذلك خلت تلك الدراسات من دراسات أجريت في مناطق فقيرة بإمكانياتها فجاءت هذه الدراسة "بنـــاء اختبـــار تكيفي لقياس القدرة الرياضية وفق الاستراتيجية ثنائية المرحلة حسب نظرية اســـتجابة الفقـــرة" لطلبة الصف الثاني الأساسي كمحاولة لقياس القدرة الرياضية في سن مبكرة و لدراسة جـــدوى فعالية الاختبارات التكيفية على تلك الشريحة من الطلبة الصغار.

الفصل الثالث

الطريقة والإجراءات

الفصل الثالث
الطريقة والإجراءات

يتناول هذا الفصل وصفا لمجتمع الدراسة وطريقة اختيار عينة الدراسة، وطريقة بناء أداة الدراسة، وإجراءات الدراسة، ووصف لكتيب الاختبار المرفق بالاختبار وإجراءات جمع البيانات والطرق الإحصائية المستخدمة في قياس قدرات الطلبة.

مجتمع الدراسة:

تألف مجتمع الدراسة من جميع طلبة الصف الثاني الأساسي في محافظة المفرق وباديتها والمنتظمين في الدراسة في الفصل الثاني وفق التقرير الإحصائي للعام الدراسي 2007/2006. والموزعين كما في الجدول (1) على ثلاثة مديريات:

الجدول (1): توزيع مجتمع الدراسة حسب المديرية والجنس

المديرية	الجنس		العدد الكلي
	ذكور	إناث	
قصبة المفرق	1245	1184	2429
البادية الشمالية الشرقية	1198	906	1204
البادية الشمالية الغربية	647	1656	2303

عينة الدراسة

من خلال الرجوع إلى المديرية الأساسية مديرية التربية والتعليم للواء قصبة المفرق تم حصر عدد المدارس في المحافظة وقامت الباحثة بزيارات لتلك المدارس لتوضيح تعليمات الاختبار للمعلمين المتعاونين والمدراء ورصد أسماء المدارس المتعاونة وعدد طلبة الصف الثاني الأساسي فيها وتم اختيار ثلاثة عينات لهذه الدراسة بالطريقة المتيسرة (Available) وهي:

العينة الأولى: من أجل التجريب الأولي

العينة الثانية: من أجل معايرة الفقرات

العينة الثالثة: من أجل القياس التكيفي

1. عينة التجريب الأولي

تألفت عينة المفحوصين الذين تقدموا للاختبار بصورته الأوليـــة مـــن (100) طالـــب وطالبه من طلبة الصف الثاني الأساسي موزعين على خمسة مدارس والجـــدول (2) يوضـــح توزيع الطلبة وفقا لمتغيري الجنس والمدرسة:

الجدول (2) توزيع أفراد عينة التجريب الأولي وفقا لمتغيري المدرسة والجنس

المدرسة	الجنس		المجموع
	ذكور	إناث	
مدرسة حي البوادي الأساسية	8	10	18
مدرسة المنارة الأساسية	7	3	10
مدرسة البشرية الأساسية	7	15	22
مدرسة الأشرفية الأساسية	11	14	25
مدرسة أم النعام الغربية الأساسية	16	9	25
المجموع	49	51	100

وقد اختيرت هذه العينة بالطريقة العشوائية من بين المدارس المتعاونة التي اختيـــرت حسب إمكانية تعاونها مع الباحثة وبذلك سميت بالعينة المتيسرة وقد استخدمت هذه العينة (عينة التجريب الأولي) في عينة تدريج الاختبار فيما بعد والتي بلغ عددها (500) طالب وطالبة.

2. عينة تدريج الفقرات

تألفت عينة تدريج الاختبار بصورته النهائية من (500) طالباً وطالبة موزعين على عدة مدارس من ضمنها (100) مفحوص جرى اختيارهم في عينة التجريب الأولي كذلك توزعــت عينة الدراسة على مديريات التربية والتعليم الثلاثة في محافظة المفــرق المتمثلــة بالمــدارس الأساسية التالية: مدرسة (الثلاج، حي البوادي، روضة بسمه، البــشرية، الــصالحية، المنــاره، رحبة ركاد ،سما السرحان، الأشرفية،أم النعام الغربية، الحباب بن المنــذر، الربيــع، شــويكة الغربية، حي الضباط) الأساسية وبلغ حجم هذه العينة (500) طالب وطالبة وهذا العدد يفوق ما أشار إليه كروكر والجينا بأن يكون عدد أفراد عينة التدريج ما بــين (5-10) أضــعاف عــدد الفقرات(Crocker and Algina, 1986) وقد أختيرت جميعها من بين المدارس التــي أبــدت رغبة في التعاون مع الباحثة لإجراء هذه الدراسة وبذلك سميت بالعينة المتيسرة والجــدول (3) يوضح توزيع الطلبة وفقا لمتغيري الجنس والمدرسة.

الجدول (3) توزيع أفراد عينة تدريج الفقرات وفقا لمتغيري الجنس والمدرسة

المدرسة	الجنس		المجموع
	ذكور	إناث	
مدرسة الثلاج الأساسية	11	5	16
مدرسة حي البوادي الأساسية	8	10	18
مدرسة روضة بسمه الأساسية	8	7	15
مدرسة البشرية الأساسية	7	15	22
مدرسة الصالحية الأساسية	12	9	21
مدرسة المنارة الأساسية	7	3	10
مدرسة رحبة ركاد الأساسية	10	17	27
مدرسة سما السرحان الأساسية	16	14	30
مدرسة الأشرفية الأساسية	11	14	25
مدرسة أم النعام الغربية الأساسية	16	9	25
مدرسة الحباب بن المنذر الأساسية	35	45	80
مدرسة الربيع الأساسية	43	47	90
مدرسة شويكة الغربية الأساسية	19	20	39
مدرسة حي الظباط الأساسية	40	42	82
المجموع	243	257	500

3.عينة القياس التكيفي

تألفت عينة القياس التكيفي من (100) طالب وطالبة تم اختيارهم من مدرسة (الــسعادة الغربية الأساسية الجديدة، مدرسة أم حبيبة الأساسية، مدرسة أم عطية الأنــصارية الأساســية) ويعزى صغر حجم عينة القياس التكيفي إلى أن الباحثة قد حرصت على تطبيق اختبارات القياس التكيفي بنفسها وذلك لان تطبيق هذه الاختبارات يحتاج إلى جهد ووقت ودقة في التطبيق وبذلك تم اختيار أفراد هذه العينة بالطريقة القصدية من العينة الأصلية للدراسة والتي ســميت بالعينــة المتيسرة ويبين الجدول (4) توزيع أفراد عينة القياس التكيفي وفقا لمتغيري المدرسة و الجنس:

الجدول (4) توزيع أفراد عينة القياس التكيفي وفقا لمتغيري المدرسة و الجنس

المدرسة	الجنس		المجموع
	ذكور	إناث	
مدرسة السعادة الغربية الأساسية	15	20	35
مدرسة أم حبيبة الأساسية	15	20	35
مدرسة أم عطية الأنصارية الأساسية	12	18	30
المجموع	42	58	100

أدوات الدراسة :

تم اعتماد كتاب الرياضيات للصف الثاني الأساسي الفصل الدراسي الثاني ودليل المعلم لمنهاج الرياضيات للمرحلة الأساسية الأولى في إعداد اختبار لقياس القدرة الرياضــية لطلبــة الصف الثاني الأساسي وقد تم اختيار وحدتي الضرب والقسمة لبنــاء الاختبــار. **وفيمــا يلــي إجراءات بناء الاختبار وفق الطريقة التقليدية .**

1. **تحليل المحتوى:** تم تحليل المحتوى لوحدتي الضرب والقسمة بكتابــة الأهــداف لمــا يتوجب على الطلبة إتقانه بعد دراسة هاتين الوحدتين وذلك بعد الإطلاع علــى كتــاب

الطالب ودليل المعلم للصف الثاني الأساسي (الفصل الدراسي الثاني) والأهداف العامة للوحدتين فيهما والملحق (1) يبين الأهداف للوحدتين السابقتين.

2. **بناء جدول المواصفات:**تم بناء جدول مواصفات بعد الإطلاع على الأهداف التدريسية المبينة في تحليل المحتوى وتصنيفها إلى المستويات الفرعية التالية: المعرفة، الاستيعاب، التطبيق، التحليل.

واستخرجت النسب المئوية لتمثيل المحتوى (وحدتي الضرب والقسمة) والأهداف وتصنيفاتها الفرعية ويبين الجدول (5) جدول المواصفات.

الجدول (5) جدول مواصفات الاختبار

المستوى / الوحدات	المعرفة %59	الاستيعاب %8	تطبيق %25	التحليل %8	المجموع %100
وحدة الضرب 65%	%38	%5	%16	%6	%65
42 فقرة	24 فقرة	3 فقرات	11 فقرة	4 فقرات	42 فقرة
وحدة القسمة35%	%21	%3	%9	%2	%35
22 فقرة	13 فقرة	2 فقرة	6 فقرات	1 فقرة	22 فقرة

3. كتابة الفقرات

تم إعداد اختبار لقياس القدرة الرياضية في وحدتي الضرب والقسمة للصف الثاني الأساسي تكون في صورته الأولية من (64) فقرة من نوع الاختيار من ثلاث بدائل اعتماداً على جدول المواصفات السابق ويتضمن الملحق (2) فقرات الاختبار بصورته الأولية بعد حذف فقرتين أقترح بعض المحكمين حذفها.

4. مراجعة وتنقيح الفقرات (صدق المحتوى)

عُرضت فقرات الاختبار بالإضافة لتحليل المحتوى وجدول المواصفات على مجموعة من المختصين في الرياضيات ومشرفي الرياضيات في مديريات التربية والتعليم في محافظة المفرق ومدرسي الجامعات والمعلمين والمدراء ممن ابدوا الرغبة في مساعدة الباحثة ومشرفي تربية للمرحلة الأساسية الأولى بهدف الحكم على تحليل المحتوى وجدول المواصفات وإبداء الملاحظات على الاختبار حيث استُخدم نموذج خاص لتقويم الاختبار يتضمنه ملحق (3) كذلك ملحق (4) الذي تضمن أسماء المحكمين للاختبار وبناء على ملاحظاتهم تم تعديل بعض فقرات الاختبار وتعديل بعض البدائل للفقرات وتحديد الزمن اللازم لتطبيق الاختبار وأصبح الاختبار بصورته الأولية مكون من (64) فقرة.

كذلك وبناءاً على ملاحظات لجنة التحكيم تم إعداد كتيب للاختبار يقدم "للفاحص" وهو المعلم الذي سيتعاون مع الباحثة في تطبيق الاختبار. أيضاً تم الاتفاق على إعداد صفحة خاصة بتعليمات الاختبار للطلبة المفحوصين وإرفاقها في مقدمة صفحات الاختبار. وقد أبدى (95%) من المحكمين موافقتهم على التوزيع المنطقي لفقرات الاختبار على القدرات الرياضية المطلوب قياسها وكذلك مناسبة جدول المواصفات لقياس التحصيل في وحدتي القسمة والضرب. وبالتالي ضمان ما يسمى بالصدق العيني أي إن عينة الفقرات ممثلة للمحتوى. وبذلك افترضت الباحثة إن الاختبار يتمتع بصدق محتوى مقبول حيث أن الباحثة قامت بتحليل منهاج الرياضيات للصف الثاني الأساسي (للفصل الدراسي الثاني) وبناء جدول مواصفات للاختبار.

كتيب الاختبار (للفاحص)

تم إعداد كتيب للمطبق أو الفاحص يقدم مع كل مجموعة من الاختبارات الفرعية التي تم تكوينها المعد لأحد الصفوف التي سيطبق عليها الاختبار لإيضاح الأمور التالية:

أ. وصف موجز للاختبار:

يتكون هذا الاختبار من فقرات لقياس القدرة الرياضية في وحدتي الضرب والقسمة التي تدرس للصف الثاني الأساسي ويمكن لمعلم الصف الاستفادة من هذا الاختبار باعتباره مقياس محكي المرجع في المهارات الواجب إتقانها وبالتالي تحديد أخطاء الطلبة كذلك لو اُعتُبر مقياس معياري المرجع بالتزود بمعايير صفية تساعد على تعيين موقع الطالب بالنسبة لباقي الطلبة في عمره ومستواه الأكاديمي للمقارنة بين الطالب وطالب آخر أو صف بصف آخر في نفس المدرسة أو مدرسة أخرى.

ب .طريقة التطبيق:

يتكون هذا الاختبار من فقرات في وحدتي الضرب والقسمة للصف الثاني الأساسي قدمت الصور مع الأسئلة التي تتطلب توضيح أكثر من شرح المعلم لفهمها من قبل الطالب ووضع رقم السؤال على يمين كل سؤال بخط غامق وعريض لتميزه وكذلك خط للفصل بين كل سؤال والسؤال الذي يليه أما البدائل فكانت ثلاثة بدائل استخدمت الأحرف (أ، ب، ج) لفصلها عن بعضها طبعت على أوراق وأرفق بها ورقة خاصة بتعليمات الاختبار وبيانات الطالب يقرأها المعلم الفاحص أمام الطلبة شفويا ويشرحها شرحا وافيا وذلك في فترة التمهيد للاختبار وذلك في الحصة الأولى من اليوم الذي عُقد فيه الاختبار.

ج. تنفيذ الاختبار:

- يشرح المعلم مثالا على السبورة على نمط أسئلة الاختبار أمام الطلبة لتعريف الطالب بكيفية الإجابة "كيف يختار الإجابة الصحيحة من بين البدائل الواردة".
- يقرأ المعلم تعليمات تنفيذ الاختبار ويشرحها.

- يقرأ المعلم كل سؤال من أسئلة الاختبار بصوت واضح ثم يتأكد من أن كل طالب أجاب عن هذا السؤال وفق معرفته مع مراعاة عدم توضيح الإجابة.
- ينتقل المعلم من سؤال إلى آخر وإعطاء وقت كافي للطلبة للإجابة عن هذا السؤال وهكذا حتى ينتهي الطلبة من إجابة جميع الأسئلة.
- يقدم الاختبار لطلبة الصف الثاني الأساسي في الحصة الثانية والحصة الثالثة ومدة كل حصة (45) دقيقة فيكون الزمن اللازم للاختبار (90) دقيقة متواصلة ولكن بعد مضي منتصف الوقت المحدد أي (40) دقيقة يعطي المعلم فترة استراحة للطلبة مدتها (10) دقائق يطلب فيها وضع الأقلام على الطاولة وممارسة نشاط لا منهجي معد مسبقاً (كاللعب بالأيدي).

د. التصحيح:

حسبت لكل فقرة أجاب عنها الطالب إجابة صحيحة علامة واحدة والتي اخطأ فيها تُعطى العلامة صفر ولا تُعطى أجزاء من العلامة أو علامات تشجيعية فقط التعزيز المعنوي بألفاظ (ممتاز، جيد، مقبول) باستخدام القلم الأحمر فقط لأن تصحيح أوراق الاختبار تم يدوياً.

و. تعليمات للمطبق (الفاحص):

- تأكد من تسليم كل طالب الأسئلة المخصصة لهذا الاختبار.
- اقرأ تعليمات التطبيق بصوت واضح.
- اشرح مثالا على السبورة قبل البدء بالاختبار وتأكد من وضوح للمثال وذلك قبل بدء وقت الاختبار.
- اقرأ كل سؤال من أسئلة الاختبار بعد تأكدك من إنهاء الطلبة للسؤال السابق.
- راعي الزمن المخصص لكل جلسة من جلسات الاختبار وفترة الاستراحة.

5. تجريب أولي للفقرات:

تم تطبيق الاختبار على عينة التجريب الأولي وبلغ عددها (100) طالب وطالبة توزعوا على خمسة مدارس وبالاعتماد على التجريب الأولي للاختبار تم حساب صعوبة الفقرات بإيجاد نسبة الطلبة الذين أجابوا على الفقرة إجابة صحيحة من بين الطلبة الذين حاولوا الإجابة عن هذه الفقرة وتم حساب القدرة التمييزية للفقرات من خلال حساب معاملات الارتباط بوينت بايسيريال (Point biserial correlation coefficient) بين الفقرة والاختبار والجدول (6) يبين معاملات الصعوبة ومعاملات التمييز لفقرات الاختبار في عينة التجريب الأولي .

الجدول (6) معاملات الصعوبة والتمييز لفقرات الاختبار لعينة التجريب الأولي

رقم الفقرة	متوسط معامل الصعوبة	معامل التمييز	رقم الفقرة	متوسط معامل الصعوبة	معامل التمييز	رقم الفقرة	معامل الصعوبة	معامل التمييز
1	0.67	0.14	23	0.59	0.62	45	0.57	0.51
2	0.60	0.44	24	0.59	0.36	46	0.37	0.62
3	0.66	0.23	25	0.56	0.44	47	0.44	0.56
4	0.58	0.42	26	0.56	0.52	48	0.43	0.61
5	0.54	0.23	27	0.55	0.46	49	0.44	0.56
6	0.56	0.24	28	0.55	0.47	50	0.46	0.45
7	0.49	0.40	29	0.91	0.18	51	0.40	0.54
8	0.56	0.33	30	0.55	0.50	52	0.36	0.44
9	0.55	0.36	31	0.60	0.53	53	0.44	0.64
10	0.47	0.47	32	0.48	0.20	54	0.36	0.55
11	0.45	0.41	33	0.45	0.45	55	0.43	0.34
12	0.59	0.51	34	0.56	0.44	56	0.48	0.65
13	0.67	0.40	35	0.44	0.54	57	0.40	0.63
14	0.51	0.46	36	0.58	0.48	58	0.41	0.59
15	0.57	0.45	37	0.36	0.33	59	0.37	0.58
16	0.55	0.38	38	0.61	0.25	60	0.35	0.50
17	0.49	0.37	39	0.39	0.44	61	0.33	0.54
18	0.43	0.41	40	0.34	0.53	62	0.33	0.46
19	0.63	0.47	41	0.46	0.28	63	0.30	0.32
20	0.60	0.38	42	0.46	0.51	64	0.38	0.43
							6.71	
21	0.42	0.55	43	0.41	0.62			
22	0.60	0.43	44	0.47	0.50			
	9.02			8.94				

يتضح من الجدول (6) أن مجموع المتوسطات الحسابية لمعاملات الصعوبة بلغ (24.67) وبلغ الانحراف المعياري (14.25) أما التباين فبلغ (2.30) وتم حساب معامل ثبات

الاتساق الداخلي للاختبار باستخدام معادلة (كودر-ريتشارد سون-20)(kr-20) وكانت قيمة معامل ثبات الاختبار (0.94) مما يدل على أن هناك تجانساً وظيفياً في الفقرات وذلك أثناء قياسها للسمة التي صُمم الاختبار لقياسها وهي "القدرة الرياضية" وهو مؤشر جيد على أحادية البعد، ولكن لم يُعتمد على هذا التحليل في حذف أية فقرات لأن الباحثة اعتمدت على برنامج (BIGSTEPS) في التحليل حسب النظرية الحديثة في القياس وبناء على نتائج تحليله حذفت بعض فقرات الاختبار.

وفي خطوة أخرى للتأكد من صدق الاختبار والتحقق من افتراض أحادية البعد تم استخدام أسلوب التحليل العاملي (Factor Analysis) باستخدام برنامج الحزمة الإحصائية للبحوث التربوية (SPSS) لتحديد البناء لعاملي للاختبار والذي يوضحه جدول (7).

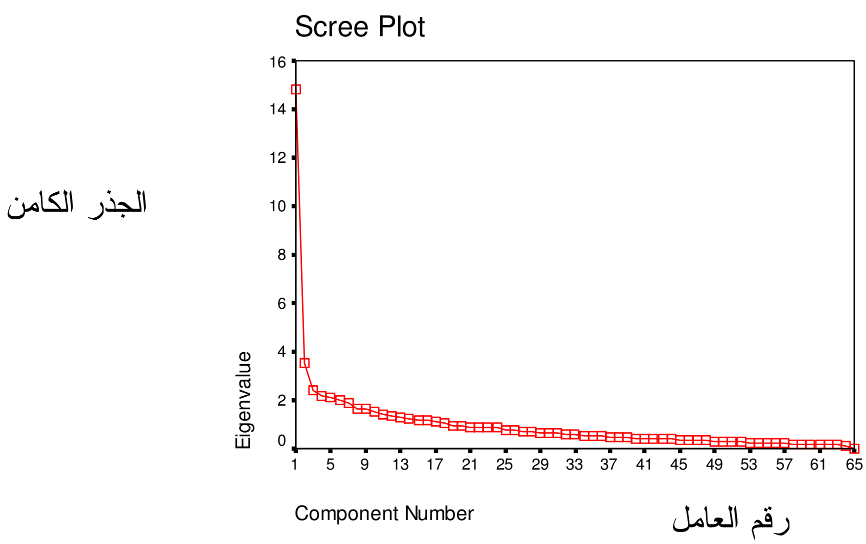
الجدول (7): قيم الجذور الكامنة للعوامل التي تشبعت بها فقرات الإختبار

Total Variance Explained

Cumulative %	% of Variance	Total	Component
التباين التراكمي المفسر	التباين المفسر	الجذر الكامل	العامل
22.797	22.797	14.818	1
28.250	5.453	3.545	2
31.921	3.671	2.386	3
35.278	3.357	2.182	4
38.516	3.238	2.105	5
41.612	3.096	2.013	6
44.493	2.881	1.872	7
47.061	2.568	1.669	8
49.553	2.492	1.620	9
51.861	2.308	1.501	10
54.072	2.211	1.437	11
56.145	2.073	1.347	12
58.124	1.979	1.287	13
60.061	1.937	1.259	14
61.885	1.823	1.185	15
63.670	1.785	1.160	16
65.348	1.678	1.091	17
66.938	1.591	1.034	18

Extraction Method: Principal Component Analysis.

ويلاحظ من الجدول (7) أن نسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني أكبر من (2) وان ما يفسره العامل الأول من التباين أكبر من(20%) وهذا مؤشر جيد على أحادية البعد كما أشار(Hambleton&Swaminathan,1985) كما يتضح تحقق افتراض أحادية البعد من خلال التمثيل البياني لقيم الجذور الكامنة للعوامل المكونة للاختبار الموضحة في الشكل (2).



الشكل (2): التمثيل البياني لقيم الجذور الكامنة للعوامل المكونة للاختبار

-افتراض الاستقلال الموضعي (Local Independence): وهذا الافتراض مكافئ لافتراض أحادية البعد (Hambleton & Swaminathan, 1985).

التطبيق الرئيسي للاختبار

طبق الاختبار على عينة بلغت (500) طالب وطالبة بتاريخ 2007/4/14 حيث وزعت أوراق الأسئلة بالإضافة لنسخة من كتيب الاختبار إلى المعلمات والمعلمين المتعاونين مع الباحثة حيث تم التنسيق مسبقا وبشكل ودي معهم للتدرب حول وقت وإجراء الاختبار ومكانه حيث يتم جمع طلبة المدرسة المعنيين بالدراسة في مكان مناسب بعد أن اعلموا بموعد الامتحان

ومكانه ويتم تطبيق الاختبار بحضور المعلم (الفاحص) ومرافقة الباحثة أحيانا لإعطاء المزيد من الأهمية للاختبار تم التطبيق في جلستين بينهما فاصل للاستراحة مدته (10) دقائق وقد طبق في الحصة الثانية والثالثة من صباح اليوم المذكور وبعد انتهاء الاختبار جمعت المعلمات نسخ الاختبار وسلمتها للباحثة في نفس اليوم.

بعد الانتهاء من عملية التطبيق على عينة الدراسة والبالغ عددها (500) طالب وطالبة تم تصحيح كراسات الاختبار يدوياً ورصد العلامة (1) للإجابة الصحيحة والعلامة (0) للإجابة الخطأ وبذلك تكون أقصى علامة لهذا الاختبار (64) ثم أستخدم برنامج (Excel) لإدخال العلامات لجهاز الحاسوب ثم نقلت إلى تطبيق (Note pad) وتم الحصول على سجل لكل طالب من الطلبة يتكون من إجاباته ثنائية التدريج (1.0) على الاختبار وكونت هذه السجلات ملفاً حاسوبياً لإجابات جميع الطلبة.

تدريج الاختبار(تقدير قيم(θ) قدرات المفحوصين وقيم (b) معامل صعوبة الفقرات)

تم تحليل استجابات أفراد عينة الدراسة الذين استجابوا عن فقرات الاختبار باستخدام برنامج (BIGSTEPS) المصمم لتحليل البيانات وفقاً لنموذج راش حيث يحذف البرنامج استجابات الأفراد الحاصلين على العلامة الكاملة أو الحاصلين على العلامة (صفر) وكذلك يحذف الفقرات التي يجيب عنها أفراد العينة جميعهم إجابة صحيحة أو تلك التي يفشل جميع الأفراد في الإجابة عليها وعلى هذا الأساس حذف البرنامج مباشرة (7) أفراد كانت علاماتهم كاملة أي حصلوا على (64) وهي العلامة الكاملة لهذا الاختبار.

تم تحليل الاستجابات للمرة الأولى وذلك لحذف الأفراد غير المطابقين للنموذج حيث تم تقدير قدرة كل فرد والخطأ المعياري في تقدير هذه القدرة وكذلك قيم إحصائي المطابقة الداخلية

والخارجية (ZSTD) وكذلك متوسط المربعات الداخلية والخارجية من (MNSQ) وتعد قـدرة الفرد غير متطابقة مع قدرات مجموعة الأفراد التي ينتمي إليها إذا كانت قيمة متوسط المربعات المناظرة لدرجته تبتعـد عـن (+1) أي تقـع خـارج حـدود المطابقـة وهـي (0.7-1.3) (Julian, 1988) وبناءً على قيم إحصائي المطابقة الخارجية للأفراد الذي يشير إلى مدى اتساق قدرة الفرد مع قدرات أفراد المجموعة التي تنتمي إليها في قياس السمة التي يقيسها الاختبار تـم استبعاد استجابات (16) فرد من الأفراد الذين استجابوا على فقرات الاختبـار لأن اسـتجاباتهم تبتعد عن توقعات النموذج والجدول (8) يوضح إحصائيات المطابقة للأفراد المستبعدين والملحق (5) يوضح إحصائيات المطابقة لجميع الأفراد كذلك ملحق (6) يظهر فيه الرسـم التوضـيحي لمطابقة الأفراد.

الجدول (8): أرقام أفراد عينة الدراسة الذين تم إستبعادهم نتيجة تجاوز إحصائي المطابقة الخارجية المقابل لقدراتهم المعيار المحدد (0.7−1.3)

رقم الفرد	متوسط المربعات الموزونة(MNSQ)	إحصائي المطابقة الخارجية (ZSTD)	رقم الفرد	متوسط المربعات الموزونة(MNSQ)	إحصائي المطابقة الخارجية (ZSTD)
26	1.69	3.7	196	1.43	1.7
225	1.64	2.2	145	1.40	1.6
271	1.53	2.0	487	1.40	1.6
98	1.48	1.7	89	1.40	3.8
498	1.56	5.4	128	0.09	-1.4
444	1.52	1.8	179	0.09	-1.4
23	1.48	4.9	254	0.09	-1.4
12	1.44	4.5	470	0.09	-1.4

يتضح من الجدول (8) أن الأفراد المستبعدين كانت متوسـطات المربعـات للمطابقـة الخارجية المقابلة لقدراتهم خارج حدود المطابقة وهذا يعني أن استجاباتهم تبتعد عن القيم التـي يتوقعها النموذج.

بعد استبعاد الأفراد الذين لم تتطابق استجاباتهم مع توقعات النموذج أعيد التحليل للمـرة الثانية لاختبار مدى ملاءمة الفقرات للنموذج فقد تم تقدير معلم الصعوبة لكل فقرة بالإضافة إلى

الخطأ المعياري في قياس الصعوبة وقيم إحصائي المطابقة الداخلية والخارجية للفقرات (ZSTD) ومتوسط المربعات للفقرات (MNSQ) وتعد صعوبة الفقرة غير متطابقة مع النموذج إذا كانت قيمة متوسط المربعات الموزونة بعيدة عن (+1) أي تقع خارج حدود المطابقة وهي (1.3-0.7) (Julian, 1988) وبناءً على قيم متوسطات المربعات الموزونة التي تشير إلى مدى اتساق صعوبة الفقرة مع صعوبة مجموعة الفقرات التي تنتمي إليها تم استبعاد الفقرات الواقعة خارج حدود المطابقة وهي (1.3-0.7) من فقرات الاختبار وبلغ عددها فقرتين والجدول (9) يبين أرقام الفقرات التي تم استبعادها من فقرات الاختبار وقيم إحصائي المطابقة لها ومتوسط المربعات الموزونة ويتضمن الملحق (7) إحصائيات المطابقة الداخلية والخارجية لجميع الفقرات كذلك الشكل (3) يظهر فيه رسم توضيحي للمطابقة.

الجدول (9) أرقام الفقرات التي تم استبعادها من فقرات الاختبار ومتوسط المربعات الموزونة وقيم إحصائي المطابقة الخارجية لها

رقم الفقرة	متوسط المربعات الموزونة (MNSQ)	إحصائي المطابقة الخارجية(ZSTD)
1	1.60	6.0
5	1.43	6.4

يلاحظ من الجدول (9) بأن هاتين الفقرتين يبتعد متوسط المربعات لهما عن (+1) وهذا مؤشر على أنها غير مطابقة وبعيدة عن توقعات النموذج.

أصبح الاختبار مكون من (62) فقرة ومن أجل الحصول على اختبارات فرعية متساوية في عدد الفقرات تم حذف فقرتين من الفقرات ذات معامل تمييز قليل تظهر في جدول (10) وكانت هاتان الفقرتان وهما الفقرة (32) حيث كان معامل تمييزها (0.11) كذلك الفقرة (3) وكان معامل تمييزها (0.14) فأصبح الاختبار مكون من (60) فقرة بصورته النهائية.

بعد استبعاد الفقرات غير المطابقة لتوقعات النموذج تم إعادة التحليل للمرة الثالثة من أجل الحصول على تقديرات نهائية متحررة لكل من صعوبة الفقرات وقدرات الأفراد.

جدول (10) ترتيب الفقرات تنازلياُ حسب قيم معلم الصعوبة وتظهر من خلاله الفقرتان المحذوفتان بسبب تدني معامل تمييزهما

ITEMS STATISTICS: MEASURE ORDER

ENTRY NUMBER	RAW SCORE	COUNT	MEASURE	ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTBIS CORR.	ITEM
63	132	477	1.12	0.11	1.21	3.4	1.22	2.1	0.23	I063
61	149	477	0.90	0.11	0.91	-1.8	0.84	-2.0	0.49	I061
62	149	477	0.90	0.11	0.96	-0.7	1.04	0.5	0.41	I062
40	155	477	0.83	0.11	0.95	-1.0	0.87	-1.6	0.48	I040
60	157	477	0.81	0.11	0.95	-1.0	0.92	-1.0	0.45	I060
52	160	477	0.77	0.11	1.07	1.3	1.05	0.6	0.36	I052
54	160	477	0.77	0.11	0.90	-1.9	0.83	-2.3	0.51	I054
37	165	477	0.71	0.11	1.18	3.5	1.18	2.2	0.25	I037
46	165	477	0.71	0.11	0.82	-3.9	0.76	-3.3	0.57	I046
59	168	477	0.68	0.11	0.85	-3.1	0.82	-2.5	0.54	I059
64	172	477	0.63	0.11	1.04	0.8	1.09	1.1	0.37	I064
39	181	477	0.53	0.11	1.02	0.4	1.03	0.4	0.40	I039
51	182	477	0.52	0.11	0.94	-1.4	0.98	-0.3	0.48	I051
57	184	477	0.50	0.11	0.79	-4.9	0.72	-4.4	0.61	I057
58	188	477	0.45	0.10	0.84	-3.7	0.79	-3.2	0.56	I058
43	190	477	0.43	0.10	0.84	-3.7	0.78	-3.4	0.57	I043
55	197	477	0.35	0.10	1.19	4.1	1.23	3.3	0.26	I055
21	198	477	0.34	0.10	0.90	-2.3	1.06	0.8	0.50	I021
48	198	477	0.34	0.10	0.83	-4.1	0.80	-3.2	0.57	I048
18	200	477	0.32	0.10	1.07	1.7	1.08	1.2	0.35	I018
49	201	477	0.31	0.10	0.91	-2.2	0.83	-2.7	0.51	I049
35	203	477	0.29	0.10	0.93	-1.6	0.95	-0.8	0.48	I035
47	203	477	0.29	0.10	0.89	-2.7	0.84	-2.7	0.52	I047
53	203	477	0.29	0.10	0.79	-5.3	0.77	-3.9	0.61	I053
33	208	477	0.24	0.10	1.03	0.7	1.13	2.0	0.39	I033
11	212	477	0.19	0.10	1.07	1.6	1.09	1.4	0.36	I011
50	213	477	0.18	0.10	1.03	0.7	1.01	0.1	0.39	I050
41	215	477	0.16	0.10	1.23	5.2	1.24	3.6	0.20	I041
42	217	477	0.14	0.10	0.94	-1.5	0.89	-1.9	0.47	I042
44	219	477	0.12	0.10	0.97	-0.7	0.96	-0.6	0.44	I044
32	220	477	0.11	0.10	1.34	7.5	1.38	5.5	0.11	I032
56	222	477	0.09	0.10	0.77	-6.0	0.71	-5.1	0.62	I056
10	224	477	0.07	0.10	0.95	-1.3	0.92	-1.2	0.46	I010
17	225	477	0.06	0.10	1.13	3.0	1.13	2.1	0.30	I017
7	226	477	0.05	0.10	1.10	2.3	1.06	0.9	0.33	I007
14	238	477	-0.08	0.10	1.00	-0.1	1.20	3.0	0.41	I014
27	258	477	-0.29	0.10	1.00	0.0	1.04	0.7	0.40	I027
9	259	477	-0.30	0.10	1.11	2.7	1.07	1.0	0.31	I009
30	259	477	-0.30	0.10	0.96	-0.9	0.90	-1.7	0.44	I030
26	261	477	-0.32	0.10	0.93	-1.9	0.90	-1.5	0.48	I026
28	262	477	-0.33	0.10	0.94	-1.6	0.97	-0.5	0.46	I028
6	263	477	-0.34	0.10	1.25	6.1	1.28	4.0	0.17	I006
8	264	477	-0.35	0.10	1.13	3.4	1.18	2.6	0.27	I008
16	264	477	-0.35	0.10	1.04	1.1	1.08	1.2	0.36	I016
25	264	477	-0.35	0.10	1.01	0.2	0.96	-0.6	0.40	I025
34	264	477	-0.35	0.10	1.02	0.4	0.96	-0.6	0.40	I034
15	268	477	-0.39	0.10	1.00	0.0	1.02	0.3	0.40	I015
45	268	477	-0.39	0.10	0.92	-2.3	0.86	-2.3	0.48	I045
36	276	477	-0.48	0.10	0.94	-1.7	0.95	-0.7	0.44	I036
23	278	477	-0.50	0.10	0.80	-5.7	0.72	-4.6	0.58	I023
4	279	477	-0.51	0.10	1.03	0.8	1.01	0.2	0.36	I004
12	281	477	-0.53	0.10	0.90	-2.7	0.91	-1.4	0.48	I012
24	284	477	-0.56	0.10	1.06	1.5	1.09	1.3	0.33	I024
31	284	477	-0.56	0.10	0.89	-3.0	0.81	-2.9	0.50	I031
2	285	477	-0.57	0.10	1.24	5.9	1.49	6.0	0.14	I002
22	287	477	-0.59	0.10	0.98	-0.4	1.18	2.4	0.39	I022
20	292	477	-0.65	0.10	1.03	0.9	1.02	0.3	0.36	I020
38	292	477	-0.65	0.10	1.19	4.7	1.26	3.3	0.20	I038
19	298	477	-0.71	0.10	0.94	-1.5	0.87	-1.8	0.44	I019
3	314	477	-0.89	0.11	1.21	4.8	1.39	4.2	0.14	I003
13	323	477	-0.99	0.11	0.97	-0.8	1.11	1.2	0.37	I013
29	437	477	-2.88	0.17	1.11	0.8	0.89	-0.4	0.20	I029
MEAN	230.	477.	0.00	0.11	1.00	-0.1	1.00	-0.1		
S.D.	54.	0.	0.63	0.01	0.13	3.0	0.17	2.5		

محذوفة

محذوفة

اختيار الفقرات لمرحلتي القياس التكيفي:

اختيار فقرات الاختبار الاستطلاعي (Routing Test)

وفق قيم معلم الصعوبة رتبت الفقرات تنازليا يوضحها جدول (10) واختيرت الفقـرات ذوات الأرقام (62,59,21,53,50,56,30,34,24,13) وعددها (10) فقرات لتـشكل اختبـار المرحلة الأولى (الاختبار الاستطلاعي) وقد تم اختيار فقرات الاختبار بحيث تغطي مـستويات الصعوبة كافة وتغطي الموضوعات الفرعية للإختبار وقد استخدم هذا الاختبار لإيجـاد تقـدير أولي لقدرة المفحوصين تمهيداً لتقديم اختبار المرحلة الثانية التكيفي.

اختيار فقرات اختبارات المرحلة الثانية

استخدمت الفقرات المدرجة حسب قيم صعوبتها لبناء خمسة اختبارات تكيفيـة ويعـود سبب بناء هذا العدد من الاختبارات أن تجمع الفقرات قد تكون من (60) فقـرة وكـل اختبـار فرعي تضمن (12) فقرة يوضحها جدول (11) الذي يبين أرقام فقرات الاختبـارات الفرعيـة ومدى صعوبة فقرات كل مجموعة.

الجدول (11) أرقام فقرات الاختبارات الفرعية ومدى صعوبة فقرات كل مجموعة (60) فقرة (5) اختبارات

الاختبار	أرقام الفقرات												مدى صعوبة الفقرات	
الاختبار الأول (12) فقرة	63	61	62	40	60	52	54	37	46	59	64	39	0.53	1.12 -
الاختبار الثاني (12) فقرة	51	57	58	43	55	21	48	18	49	35	47	53	0.29	0.52 -
الاختبار الثالث (12) فقرة	33	11	50	41	42	44	56	10	17	7	14	27	0.29-	0.24 -
الاختبار الرابع (12) فقرة	9	30	26	28	6	8	16	25	34	15	45	36	0.48-	0.3- -
الاختبار الخامس (12) فقرة	23	4	12	24	31	2	22	20	38	19	13	29	2.88-	0.5- -

ملاحظة :يجري تحديد الفقرات التي ستقدم للمفحوصين بعد مرحلة الاختبار الاستطلاعي اعتماد على قيمة معلم القدرة المتحققة من الاختبار الاستطلاعي .

تطبيق اختبارات القياس التكيفي

تطبيق الاختبار الاستطلاعي:

تم اختبار(100)طالب من طلبـــة" مدرســـة الـــسعادة الغربيـــة المختلطـــة الأساســـية الجديـــدة، مدرســـة أم حبيبـــة الأساســـية، مدرســـة أم عطيـــة الأنـــصارية الأساســـية " لإجـــراء القيـــاس التكيفـــي علـــيهم مـــنهم (42) طالبـــا و(58) طالبـــة توزعـــوا علـــى (3) مدارس وسبب اختيار الطلبة من تلك المـــدارس هـــو عمـــل الباحثـــة بـــالقرب منهـــا كـــي تتمكن من متابعة الاختبار التكيفي بنفسها بالتنقل بين طلبـــة الـــصف الثـــاني الأساســـي بعـــد ذلـــك جـــرى تحليـــل لبيانـــات الاختبـــار الاســـتطلاعي بإدخـــال اســـتجابات الطلبـــة علـــى الاختبـــار الاســـتطلاعي إلـــى الحاســـوب واســـتخدام برنـــامج (BIGSTEPS) لتقدير قدرة المفحوصين (θ_1) والخطأ المعياري لتقـــدير القـــدرة وذلـــك كمـــا هـــو موضـــح في الجدول (12).

الجدول (12): نتائج تحليل الاختبار الاستطلاعي تقدير(θ_1) و(SEE_1)

SEE_1	θ_1	تسلسل	SEE_1	θ_1	تسلسل	SEE_1	θ_1	تسلسل
0.66	0.42	060	0.70	0.88	020	1.86	3.55	005
0.66	0.42	067	0.70	0.88	024	1.86	3.55	008
0.66	0.42	072	0.70	0.88	026	1.86	3.55	018
0.66	0.42	073	0.70	0.88	027	1.86	3.55	033
0.66	0.42	075	0.70	0.88	029	1.86	3.55	038
0.66	0.42	085	0.70	0.88	031	1.86	3.55	055
0.66	0.42	093	0.70	0.88	035	1.86	3.55	063
0.66	0.42	095	0.70	0.88	036	1.86	3.55	068
0.64	0.00	011	0.70	0.88	040	1.86	3.55	090
0.64	0.00	015	0.70	0.88	049	1.6	2.26	001
0.64	0.00	019	0.70	0.88	056	1.6	2.26	007
0.64	0.00	021	0.70	0.88	057	1.6	2.26	009
0.64	0.00	022	0.70	0.88	059	1.6	2.26	028
0.64	0.00	039	0.70	0.88	061	1.6	2.26	032
0.64	0.00	041	0.70	0.88	065	1.6	2.26	034
0.64	0.00	045	0.70	0.88	066	1.6	2.26	051
0.64	0.00	046	0.70	0.88	070	1.6	2.26	058
0.64	0.00	047	0.70	0.88	074	1.6	2.26	062
0.64	0.00	069	0.70	0.88	081	1.6	2.26	064
0.64	0.00	071	0.70	0.88	082	1.6	2.26	083
0.64	0.00	067	0.70	0.88	084	1.6	2.26	086
0.64	0.00	080	0.70	0.88	094	1.6	2.26	098
0.64	0.00	091	0.70	0.88	096	0.80	1.43	002
0.64	0.00	092	0.70	0.88	097	0.80	1.43	004
0.66	-0.42	012	0.70	0.88	099	0.80	1.43	013
0.66	-0.42	014	0.70	0.88	100	0.80	1.43	043
0.66	-0.42	042	0.66	0.42	017	0.80	1.43	052
0.66	-0.42	044	0.66	0.42	023	0.80	1.43	054
0.66	-0.42	077	0.66	0.42	025	0.80	1.43	078
0.66	-0.42	079	0.66	0.42	030	0.80	1.43	087
0.66	-0.42	089	0.66	0.42	037	0.70	0.88	006
0.70	-0.88	003	0.66	0.42	048	0.70	0.88	010
0.70	-0.88	053	0.66	0.42	050	0.70	0.88	016
0.70	-0.88	088						

يلاحظ من الجدول (12) ان تقدير قدرة الأفراد بواسطة الاختبار الاستطلاعي تراوحت بين (3.55) لوجيت كحد أعلى و(0.88−) لوجيت كحد أدنى، والبرنامج المستخدم (BIGSTEPS) يعتمد على تقدير القدرة (θ_1) في ترتيب القيم للقدرة تنازلياً ولذلك تظهر في الجدول قيم (θ_1) مرتبة تنازلياً.

تطبيق اختبارات المرحلة الثانية من القياس التكيفي :

استخدم تقدير قدرة المفحوصين في مرحلة القياس التكيفي الأولـــى مرحلـــة الاختبـــار الاستطلاعي لتوجيه المفحوصين إلى واحد من اختبارات المرحلة الثانية حيث

قدم الاختبار الأول للطلبة الذين تراوحت قدراتهم بين (3.55) و(2.26).

قدم الاختبار الثاني للطلبة الذين تراوحت قدراتهم بين (2.26) و(0.88).

قدم الاختبار الثالث للطلبة الذين تراوحت قدراتهم بين (0.88) و(0.42).

قدم الاختبار الرابع للطلبة الذين تراوحت قدراتهم بين (0.42) و(0.00).

قدم الاختبار الخامس للطلبة الذين تراوحت قدراتهم بين (0.00) و(0.88-).

بعد تفريغ بيانات استجابات الطلبة على اختبارات القيـــاس التكيفـــي اســـتخدم برنـــامج (BIGSTEPS) لتقدير قدرة أفراد عينة القياس التكيفي والخطأ المعياري في تقدير قدرتهم وذلك بحساب (θ_2) خلال استجابات هؤلاء الطلبة على فقرات الاختبار الاستطلاعي واختبار القيـــاس في المرحلة الثانية.

تطبيق فقرات الاختبار جميعها على عينة القياس التكيفي:

بهدف المقارنة بين تقدير القدرة باستخدام القياس التكيفي وتقـــديرها باســـتخدام فقـــرات الاختبار جميعها طبق الاختبار بجميع فقراته الـــ (60) على عينة القياس التكيفي وأدخلت بيانات استجابات أفراد العينة إلى الحاسوب وتم حساب قدرة كل منهم (θ_3) والخطـــأ المعيـــاري فـــي التقدير.

مقارنة قيم القدرة بمحكات خارجية:

بهدف فحص فاعلية القياس التكيفي تم جمع علامات الطلبة المدرسية التي حصلوا عليها من خلال ما أعده المعلمون من اختبارات يومية وشهرية لوحدتي الضرب والقسمة وتم حـــساب معامل ارتباط بيرسون بين قيم (θ) والعلامات المدرسية.

الفصل الرابع
النتائج

الفصل الرابع
النتائج

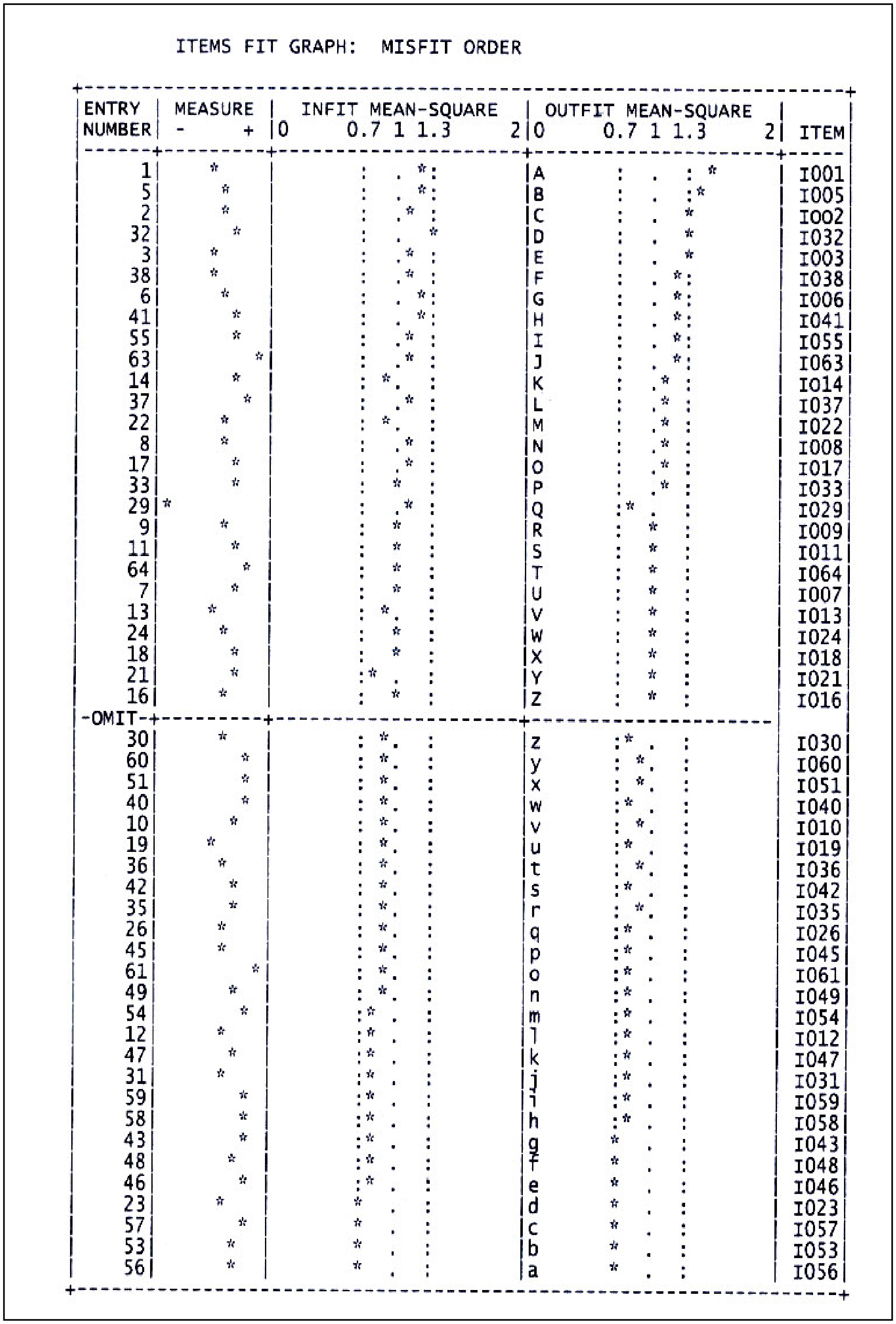
يتناول هذا الفصل عرضا لنتائج الإجابة عن الأسئلة المتعلقة بفاعلية القياس التكيفي في تقويم التحصيل في مبحث الرياضيات للصف الثاني الأساسي من خلال الإجابة على أسئلة بحثية أعدت من وحدتين في مبحث الرياضيات (الضرب والقسمة) للصف الثاني الأساسي وستعرض في هذا الفصل نتائج الإجابة على الأسئلة البحثية التي وردت في الفصل الأول والتي تم التوصل إليها من خلال البيانات التي تم جمعها من تطبيق الاختبارات التكيفية على عينة الدراسة.

أولا: النتائج المتعلقة بالسؤال الأول:

ما تقديرات الصعوبة لكل فقرة من فقرات الاختبار في عينة التدريج ؟

تم بناء اختبار تحصيلي مكون من (64) فقرة بصورته الأولية وطبق هذا الاختيار على عينة تجريبية مكونة من (100) طالبا وطالبة وذلك بهدف تجريب فقرات الاختبار حيث حللت نتائج التطبيق باستخدام النظرية الكلاسيكية فتم حساب صعوبة الفقرات والقدرة التميزية للفقرات وتراوحت معاملات الصعوبة بين (0.91) و(0.30) وتراوحت معاملات التميز بين (0.65) و(0.14) وكذلك تم حساب معامل ثبات الاتساق الداخلي للاختبار باستخدام (KR-20) وكانت قيمته (0.94) واستخدم برنامج الحزمة الإحصائية للبحوث التربوية (SPSS) لتحديد البناء العاملي للاختبار ثم طبق الاختبار على عينة بلغت (500) طالب وطالبة وصححت كراسات الاختبار يدوياً ورصدت العلامة (1) للإجابة الصحيحة والعلامة (0) للإجابة الخاطئة ثم استخدم برنامج (EXCEL) لإدخال الاستجابات لجهاز الحاسوب ونقلت لتطبيق (Notepad) وتم تحليل استجابات أفراد عينة الدراسة الذين استجابوا عن فقرات الاختيار باستخدام برنامج (BIGSTEPS) المصمم لتحليل البيانات وفقاً لنموذج راش حيث يحذف

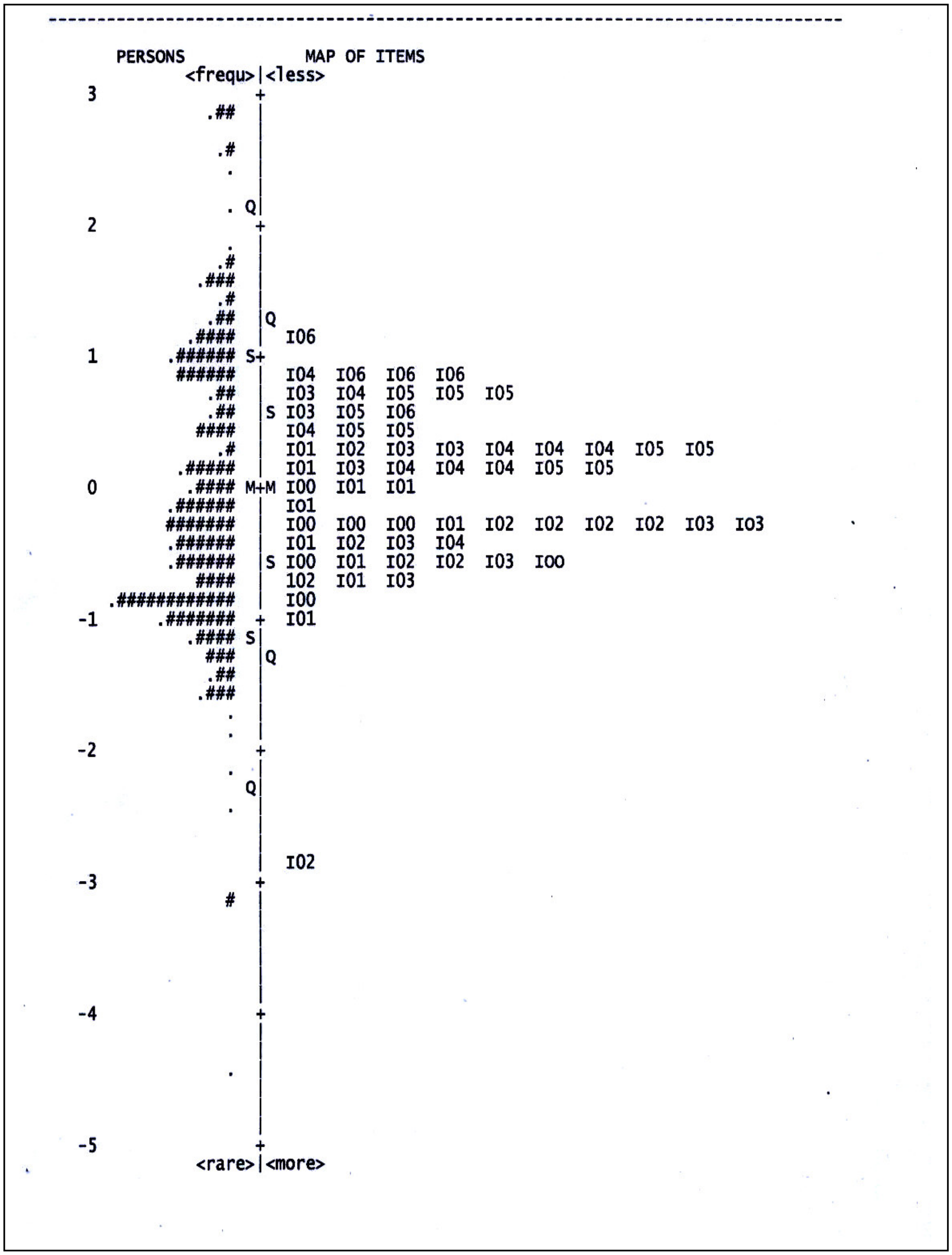
البرنامج استجابات الأفراد الحاصلين على العلامة الكاملة (64) أو الحاصلين على العلامة (0) ولذلك حذف البرنامج مباشرة (7) أفراد كانت علاماتهم كاملة، وتم تحليل الاستجابات للمرة الأولى واستخدمت بيانات المطابقة الداخلية والمطابقة الخارجية (in fit and out fit) للأفراد لحذف (16) فرد ممن لا يتطابقون مع النموذج المستخدم وذلك لان المفحوصين غير المطابقين لافتراضات النموذج قد يشكلون سبباً في عدم مطابقة بعض الفقرات للنموذج لذا يجب إعادة التحليل بعد حذف هؤلاء الأفراد وقبل التحقق من مطابقة الفقرات للنموذج ويمتاز برنامج (BIGSTEPS) بقدرته على تحديد الأفراد غير المطابقين للنموذج من خلال تقديره للمطابقة الداخلية والخارجية للأفراد ثم أعيد التحليل بعد ذلك وتم تحديد الفقرات غير المطابقة للنموذج حيث تم حذف الفقرتين (1,5) لتواجدهن خارج حدود المطابقة كذلك حذفت الفقرتين (3,32) من اجل الحصول على اختبارات فرعية متساوية في عدد الفقرات حيث كان معامل التمييز لهاتين الفقرتين هو الأقل من بين الفقرات وفي الشكل (3) يظهر الرسم التوضيحي لمطابقة الفقرات



الشكل (3): رسم توضيحي للمطابقة (جميع الفقرات)

يظهر في الشكل أن جميع الفقرات قد وقعت ضمن الحدود المقبولة للمطابقة الداخلية والخارجية باستثناء الفقرات (5، 1) كذلك الفقرتين (3، 32) تظهر على أدنى أطـراف الحـدود المقبولـة للمطابقة المعتمدة في برنامج (BIGSTEPS).

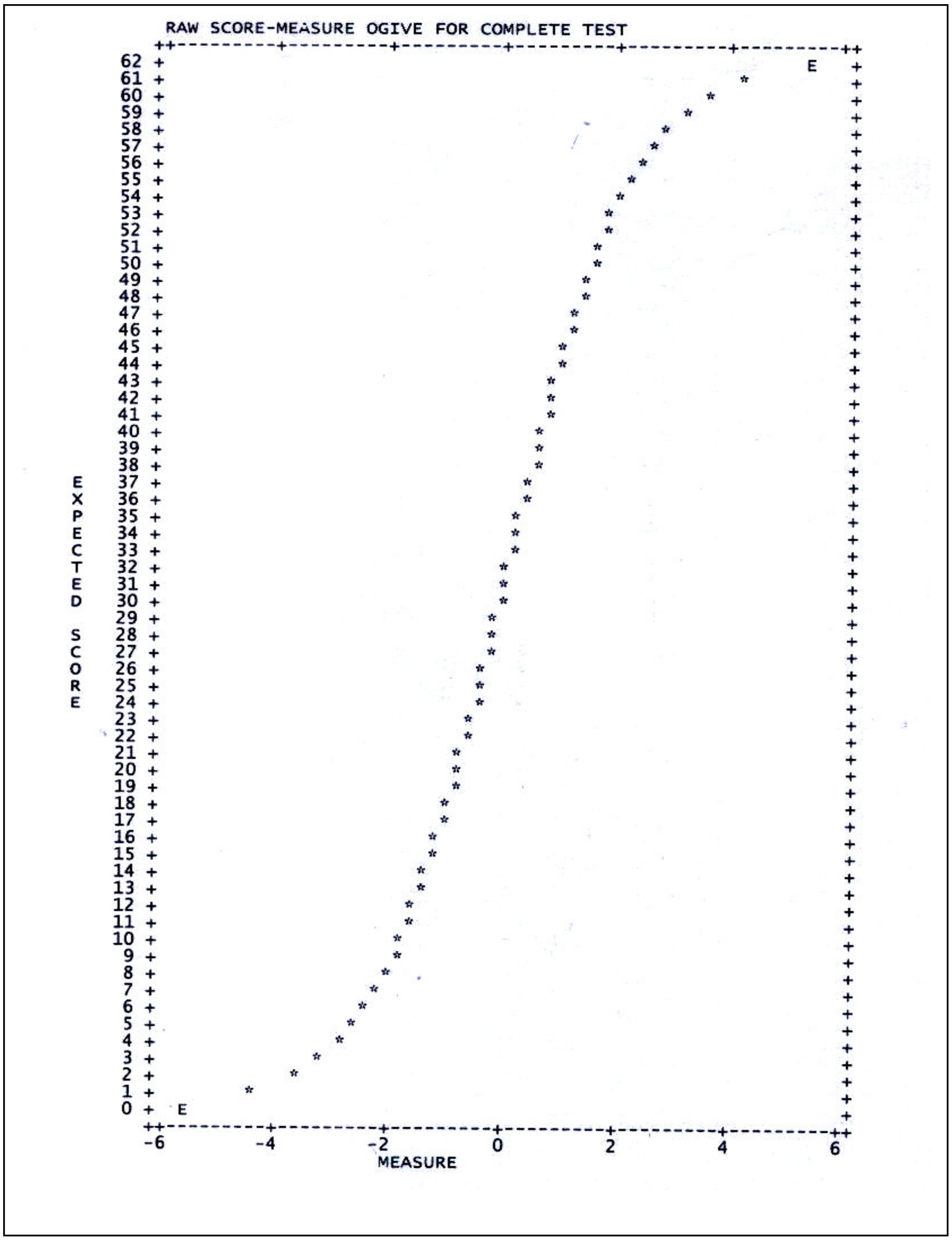
ثم أعيد التحليل حيث تم تدريج الفقرات بتقدير صعوبة كل فقرة والخطـأ المعيـاري للتقدير والموضحة سابقاً في جدول (10) حيث يظهر من الجدول أن صعوبة الفقرات تراوحت بين (1.12) لوجيت و(2.88-) لوجيت والخطأ المعياري في تقدير معلم الصعوبة تراوح بـين (0.11) لوجيت و(0.17) لوجيت ويفترض برنامج (BIGSTEPS) بان الفقرات التي تقع مـا بين (0.7 , 1.3) تحقق المطابقة الداخلية والخارجية لنموذج راش ويتأثر مفهوم المطابقة الداخلية (in fit) بنمط استجابة المفحوص حيث يفترض بان يجيب المفحوص عن الفقرات الأسـهل ولا يجيب عن الفقرات الأكثر صعوبة فإذا حدث العكس فان معيار المطابقة الداخلية يتأثر ويصبح المفحوص أو الفقرة موضوع الاهتمام غير مطابقة داخليا للنمـوذج أمـا المطابقـة الخارجيـة (out fit) وهي الأقل تهديداً للقياس فتتأثر بالتطرف في استجابات المفحوصين على الفقرة مثل التخمين الذكي أو اللامبالاة أو الخطأ في طريقة الإجابة أو إتباع نمطيـة معينـة فـي طريقـة الإجابات (Lincare, 2000).



شكل (4) توزيع الأفراد حسب قدرتهم والفقرات حسب صعوبتها

من الشكل (4) يظهر أن معظم الفقرات كانت ضمن مدى قدرة الأفــراد فقــد تركــزت قدراتهم في وسط التدريج وتجمعت كذلك معظم فقرات الاختبار في وسط التدريج.

أما الشكل (5) فيمثل منحنى خصائص الاختبار الذي يمكن من خلالـــه تحديـــد القـــدرة المناظرة لعدد الإجابات الصحيحة ومنه يلاحظ بان الزيادة في عدد الإجابات الـــصحيحة يقابلـــه زيادة في مقدار القدرة المقيسة.



شكل (5): منحنى خصائص الاختبار

ثانيا : النتائج المتعلقة بالسؤال الثاني :

ما تقديرات القدرة للأفراد (θ_1، θ_2، θ_3) التي تم تقديرها باستخدام كل من الاختبار الاستطلاعي، الاختبار التكيفي الثنائي، جميع فقرات الاختبار؟

أ- استخدم الاختبار الاستطلاعي لتقدير قدرة عينة القياس التكيفي (θ_1) والجدول (12) السابق يبين تقديرات قدرة المفحوصين باللوجيت ومقدار الخطأ المعياري في التقدير وقد بلغ متوسط تقدير القدرة للمفحوصين (0.74) لوجيت وبانحراف معياري(0.83) ومتوسط الخطأ المعياري للتقدير (0.74) لوجيت وبانحراف معياري (0.14).

ب- استخدمت تقديرات القدرة في الاختبار الاستطلاعي(θ_1) لتوزيع المفحوصين على اختبارات المرحلة الثانية والجدول (13) يوضح توزيع المفحوصين على اختبارات المرحلة الثانية .

الجدول (13) يوضح توزيع المفحوصين على اختبار المرحلة الثانية

رقم الفرد	θ_1	الاختبار المناسب	رقم الفرد	θ_1	الاختبار المناسب	رقم الفرد	θ_1	الاختبار المناسب
1	2.26	الثاني	35	0.88	الثالث	69	0.00	الرابع
2	1.43	الثاني	36	0.88	الثالث	70	0.88	الثاني
3	-0.88	الخامس	37	0.42	الثالث	71	0.00	الرابع
4	1.43	الثاني	38	3.55	الأول	72	0.42	الثالث
5	3.55	الأول	39	0.00	الرابع	73	0.42	الثالث
6	0.88	الثالث	40	0.88	الثالث	74	0.88	الثاني
7	2.26	الثاني	41	0.00	الرابع	75	0.42	الثالث
8	3.55	الأول	42	-0.42	الخامس	76	0.00	الثالث
9	2.26	الثاني	43	1.43	الثاني	77	-0.42	الثالث
10	0.88	الثالث	44	-0.42	الخامس	78	1.43	الثاني
11	0.00	الرابع	45	0.00	الرابع	79	-0.42	الخامس
12	-0.42	الخامس	46	0.00	الرابع	80	0.00	الرابع
13	1.43	الثاني	47	0.00	الرابع	81	0.88	الثالث
14	-0.42	الخامس	48	0.42	الثالث	82	0.88	الثالث
15	0.00	الثالث	49	0.88	الثالث	83	2.26	الثاني
16	0.88	الثالث	50	0.42	الثالث	84	0.88	الثاني
17	0.42	الرابع	51	2.26	الأول	85	0.42	الرابع
18	3.55	الأول	52	1.43	الثاني	86	2.26	الثاني
19	0.00	الرابع	53	-0.88	الخامس	87	1.43	الثاني
20	0.88	الثالث	54	1.43	الثاني	88	-0.88	الخامس
21	0.00	الرابع	55	3.55	الأول	89	-0.42	الخامس
22	0.00	الرابع	56	0.88	الثالث	90	3.55	الأول
23	0.42	الرابع	57	0.88	الثالث	91	0.00	الرابع
24	0.88	الثالث	58	2.62	الثالث	92	0.00	الرابع
25	0.42	الرابع	59	0.88	الثالث	93	0.42	الثالث
26	0.88	الثالث	60	0.42	الرابع	94	0.88	الثاني
27	0.88	الثالث	61	0.88	الثاني	95	0.42	الثالث
28	2.26	الأول	62	2.26	الثاني	96	0.88	الثاني
29	0.88	الثالث	63	3.55	الأول	97	0.88	الثالث
30	0.42	الرابع	64	2.26	الثاني	98	2.26	الثاني
31	0.88	الثالث	65	0.88	الثاني	99	0.88	الثاني
32	2.26	الثاني	66	0.88	الثاني	100	0.88	الثاني
33	3.55	الأول	67	0.42	الثالث			
34	2.26	الثاني	68	3.55	الأول			

وبعد جمع البيانات عن استجابات المفحوصين على هذه الاختبارات تم إدخالها للحاسوب وإيجاد تقدير جديد لقدرة المفحوصين (θ_2) يعتمد على استجاباتهم على فقرات الاختبار الاستطلاعي وفقرات اختبار المرحلة الثانية التكيفي والجدول (14) يبين تقديرات قدرة المفحوصين (θ_2) والأخطاء المعيارية في القياس.

جدول (14) تقديرات قدرة المفحوصين والأخطاء المعيارية في تقديرها باستخدام فقرات الاختبار الاستطلاعي وفقرات المرحلة الثانية من القياس التكيفي (θ_2) (SEE_2)

رقم الفرد	θ_2	SEE_2	رقم الفرد	θ_2	SEE_2	رقم الفرد	θ_2	SEE_2
033	3.70	1.85	094	1.12	0.67	041	0.34	0.59
005	2.44	1.05	096	1.12	0.67	045	0.34	0.59
008	2.44	1.05	006	0.71	0.62	047	0.34	0.59
018	2.44	1.05	010	0.71	0.62	057	0.34	0.59
032	2.44	1.05	016	0.71	0.62	059	0.34	0.59
038	2.44	1.05	017	0.71	0.62	060	0.34	0.59
051	2.44	1.05	020	0.71	0.62	067	0.34	0.59
063	2.44	1.05	024	0.71	0.62	069	0.34	0.59
064	2.44	1.05	026	0.71	0.62	071	0.34	0.59
090	2.44	1.05	029	0.71	0.62	073	0.34	0.59
098	2.44	1.05	031	0.71	0.62	080	0.34	0.59
001	1.64	0.78	035	0.71	0.62	081	0.34	0.59
007	1.64	0.78	036	0.71	0.62	082	0.34	0.59
009	1.64	0.78	043	0.71	0.62	092	0.34	0.59
028	1.64	0.78	048	0.71	0.62	095	0.34	0.59
034	1.64	0.78	049	0.71	0.62	012	0.00	0.58
054	1.64	0.78	050	0.71	0.62	014	0.00	0.58
055	1.64	0.78	070	0.71	0.62	021	0.00	0.58
058	1.64	0.78	074	0.71	0.62	025	0.00	0.58
062	1.64	0.78	078	0.71	0.62	042	0.00	0.58
068	1.64	0.78	084	0.71	0.62	044	0.00	0.58
086	1.64	0.78	085	0.71	0.62	072	0.00	0.58
087	1.64	0.78	093	0.71	0.62	075	0.00	0.58
099	1.64	0.78	097	0.71	0.62	076	0.00	0.58
100	1.47	0.79	011	0.34	0.59	091	0.00	0.58
002	1.12	0.67	015	0.34	0.59	003	-0.34	0.59
004	1.12	0.67	019	0.34	0.59	046	-0.34	0.59
013	1.12	0.67	022	0.34	0.59	053	-0.34	0.59
052	1.12	0.67	023	0.34	0.59	077	-0.34	0.59
056	1.12	0.67	027	0.34	0.59	079	-0.34	0.59
061	1.12	0.67	030	0.34	0.59	088	-0.34	0.59
065	1.12	0.67	037	0.34	0.59	086	-0.34	0.59
066	1.12	0.67	039	0.34	0.59			
083	1.12	0.67	040	0.34	0.59			

يلاحظ من الجدول (14) أن تقدير القدرة باستخدام الاختبار الاستطلاعي وفقرات المرحلة الثانية من القياس التكيفي (θ_2) قد تراوحت بين (3.70) لوجيت كحد أعلى و(0.34−) لوجيت كحد أدنى، كذلك الخطأ المعياري تراوح بين (1.85) و(0.59).

ج- كانت المرحلة الأخيرة في هذه الدراسة هي تعريض أفراد عينة القياس التكيفي إلى جميع فقرات الاختبار والبالغ عددها (60) فقرة، وجمعت استجابتهم على هذه الفقرات، وتم تحليلها، والجدول (15) يوضح تقدير قدرة المفحوصين (θ_3) والأخطاء المعيارية في التقدير.

الجدول (15) تقدير قدرة المفحوصين باستخدام جميع فقرات الاختبار θ_3

رقم الفرد	θ_3	SEE_3	رقم الفرد	θ_3	SEE_3	رقم الفرد	θ_3	SEE_3
064	5.53	1.83	026	2.06	0.39	057	1.24	0.31
098	5.53	1.83	028	2.06	0.39	060	1.24	0.31
033	4.32	1.01	036	2.06	0.39	067	1.24	0.31
052	4.32	1.01	041	2.06	0.39	069	1.24	0.31
058	4.32	1.01	061	2.06	0.39	094	1.24	0.31
063	4.32	1.01	065	2.06	0.39	021	0.15	0.30
001	3.60	0.72	097	2.06	0.39	030	1.06	0.30
032	3.60	0.72	003	1.92	0.37	031	1.06	0.30
051	3.60	0.72	007	1.92	0.37	074	1.06	0.30
062	3.60	0.72	008	1.92	0.37	044	0.97	0.29
086	3.60	0.72	014	1.92	0.37	049	0.97	0.29
087	3.60	0.72	016	1.92	0.37	082	0.97	0.29
090	3.60	0.72	017	1.92	0.37	073	0.88	0.29
099	3.60	0.72	025	1.92	0.37	081	0.88	0.29
100	3.60	0.72	040	1.92	0.37	020	0.80	0.29
002	2.86	0.52	006	1.79	0.36	045	0.80	0.29
022	2.86	0.52	027	1.79	0.36	048	0.80	0.29
024	2.86	0.52	035	1.79	0.36	050	0.80	0.29
005	2.61	0.47	039	1.79	0.36	072	0.80	0.29
011	2.61	0.47	043	1.79	0.36	093	0.80	0.29
013	2.61	0.47	056	1.79	0.36	080	0.64	0.28
024	2.61	0.47	078	1.79	0.36	076	0.56	0.28
038	2.61	0.47	084	1.79	0.36	092	0.56	0.28
065	2.61	0.47	096	1.79	0.36	075	0.40	0.28
068	2.61	0.47	009	1.66	0.34	046	0.33	0.28
083	2.61	0.47	070	1.66	0.34	047	0.33	0.28
004	2.40	0.44	029	0.55	0.33	077	0.18	0.27
023	2.40	0.44	037	0.55	0.33	091	0.18	0.27
034	2.40	0.44	042	0.55	0.33	053	0.02	0.27
055	2.40	0.44	059	0.55	0.33	079	-0.13	0.28
066	2.40	0.44	071	0.55	0.33	088	-0.36	0.28
012	2.22	0.41	095	0.55	0.33	089	-0.44	0.28
015	2.06	0.39	010	1.34	0.32			
018	2.06	0.39	019	1.34	0.32			

يلاحظ من الجدول أن تقديرات القدرة باستخدام جميع فقرات الاختبار (θ_3) قد تراوحت بين (5.53) لوجيت كحد أعلى و (0.44-) لوجيت كحد أدنى كذلك تراوحت الأخطاء المعيارية بين (1.83)و(0.28).

ثالثا : النتائج المتعلقة بالسؤال الثالث:

ما درجة اقتراب قيمة تقديرات القدرة في مرحلتي القياس التكيف (θ_1، θ_2) من تقدير قيمة القدرة(θ_3) عند تطبيق جميع فقرات الاختبار. ومن خلال الجدول (16) تظهر مقارنة بين تقديرات قدرة المفحوصين عبر ثلاثة مراحل من تقدير القدرة بعد إعادة ترتيب لتسلسل الأفراد لان البرنامج المستخدم يعتمد على تقدير القدرة في ترتيب التسلسل لذلك قامت الباحثة بإعادة ترتيب تسلسل الأفراد من اجل المقارنة بين تقديرات القدرة لنفس المفحوص .

الجدول (16) تقدير قدرة المفحوصين عبر مراحل القياس المختلفة .

رقم الفرد	θ_1	θ_2	θ_3	رقم الفرد	θ_1	θ_2	θ_3	رقم الفرد	θ_1	θ_2	θ_3
1	2.26	1.64	3.60	35	0.88	0.71	1.79	69	0.00	0.34	1.24
2	1.43	1.12	2.86	36	0.88	0.71	2.06	70	0.88	0.71	1.66
3	-0.88	-0.34	1.92	37	0.42	0.34	1.55	71	0.00	0.34	1.55
4	1.43	1.12	2.40	38	3.55	2.44	2.61	72	0.42	0.00	0.80
5	3.55	2.44	2.61	39	0.00	0.34	1.79	73	0.42	0.34	0.88
6	0.88	0.71	1.79	40	0.88	0.34	1.92	74	0.88	0.71	1.06
7	2.26	1.64	1.92	41	0.00	0.34	2.06	75	0.42	0.00	0.40
8	3.55	2.44	1.92	42	-0.42	0.00	1.55	76	0.00	0.00	0.56
9	2.26	1.64	1.66	43	1.43	0.71	1.79	77	-0.42	-0.34	0.18
10	0.88	0.71	1.34	44	-0.42	0.00	0.97	78	1.43	0.71	1.79
11	0.00	0.34	2.61	45	0.00	0.34	0.80	79	-0.42	-0.34	0.13
12	-0.42	0.00	2.22	46	0.00	-0.34	0.33	80	0.00	0.34	0.64
13	1.43	1.12	2.61	47	0.00	0.34	0.33	81	0.88	0.34	0.88
14	-0.42	0.00	1.92	48	0.42	0.71	0.80	82	0.88	0.34	0.97
15	0.00	0.34	2.06	49	0.88	0.71	0.97	83	2.26	1.12	2.61
16	0.88	0.71	1.92	50	0.42	0.71	0.80	84	0.88	0.71	1.79
17	0.42	0.71	1.92	51	2.26	2.44	0.60	85	0.42	0.71	2.06
18	3.55	2.44	2.06	52	1.43	1.12	4.32	86	2.26	1.64	3.60
19	0.00	0.34	1.34	53	-0.88	-0.34	0.02	87	1.43	1.64	3.60
20	0.88	0.71	0.80	54	1.43	1.64	2.86	88	-0.88	-0.34	-0.36
21	0.00	0.00	1.15	55	3.55	1.64	2.40	89	-0.42	-0.34	-0.44
22	0.00	0.34	2.86	56	0.88	1.12	1.79	90	3.55	2.44	3.60
23	0.42	0.34	2.40	57	0.88	0.34	1.24	91	0.00	0.00	0.18
24	0.88	0.71	2.61	58	2.26	1.64	4.32	92	0.00	0.34	0.56
25	0.42	0.00	1.92	59	0.88	0.34	1.55	93	0.42	0.71	0.80
26	0.88	0.71	2.06	60	0.42	0.34	1.24	94	0.88	1.12	1.24
27	0.88	0.34	1.79	61	0.88	1.12	2.06	95	0.42	0.34	1.55
28	2.26	1.64	2.06	62	2.26	1.64	3.60	96	0.88	1.12	1.79
29	0.88	0.71	1.55	63	3.55	2.44	4.32	97	0.88	0.71	2.06
30	0.42	0.34	1.06	64	2.26	2.44	5.53	98	2.26	2.44	5.53
31	0.88	0.71	1.06	65	0.88	1.12	2.61	99	0.88	1.64	3.60
32	2.26	2.44	3.60	66	0.88	1.12	2.40	100	0.88	1.47	3.60
33	3.55	3.70	4.32	67	0.42	0.34	1.24				
34	2.26	1.64	2.40	68	3.55	1.64	2.61				

حيث أن θ_1 : تقدير القدرة باستخدام فقرات الاختبار الاستطلاعي.

θ_2 : تقدير القدرة باستخدام فقرات الاختبار الاستطلاعي واختبار المرحلة الثانية.

θ_3 : تقدير القدرة باستخدام جميع فقرات الاختبار.

ويلاحظ من النتائج بأن تقدير القدرة قد تباين باستخدام كل من الاختبار الاستطلاعي، الاختبار التكيفي الثنائي، جميع فقرات الاختبار فقد بلغ متوسط تقدير القدرة عند تطبيق الاختبار الاستطلاعي (0.74) لوجيت ومتوسط الخطأ المعياري لتلك القدرة بلغ (0.74) لوجيت وبلغ متوسط تقدير القدرة باستخدام الاختبار التكيفي الثنائي (0.82) لوجيت بمتوسط خطأ معياري بلغ (0.68) لوجيت وبلغ متوسط تقدير القدرة باستخدام فقرات الاختبار التكيفي جميعها (1.85) لوجيت بمتوسط خطأ معياري(0.42) لوجيت. يلاحظ أيضا في جدول (16) أن "توافيق" القدرات كانت (3×2)=6 أي ستة حالات مثلت النسب كما يلي:

نسبة $\theta_3 < \theta_2 < \theta_1$ وبلغت (4%)

نسبة $\theta_1 < \theta_3 < \theta_2$ وبلغت (2%)

نسبة $\theta_2 < \theta_1 < \theta_3$ وبلغت (45%)

نسبة $\theta_1 < \theta_2 < \theta_3$ وبلغت (40%)

نسبة $\theta_2 < \theta_3 < \theta_1$ وبلغت (7%)

نسبة $\theta_3 < \theta_1 < \theta_2$ وبلغت (2%)

وفي جدول (17) تظهر الأخطاء المعيارية في تقدير قدرات المفحوصين عبر مراحل القياس المختلفة.

الجدول (17) الأخطاء المعيارية في تقدير قدرة المفحوصين عبر مراحل القياس المختلفة

رقم الفرد	SEE_1	SEE_2	SEE_3	رقم الفرد	SEE_1	SEE_2	SEE_3	رقم الفرد	SEE_1	SEE_2	SEE_3
1	1.6	0.78	0.72	34	1.6	0.67	0.44	67	0.66	0.59	0.31
2	0.80	0.67	0.52	35	0.76	0.62	0.36	68	1.86	0.78	0.47
3	0.70	0.59	0.37	36	0.70	0.62	0.39	69	0.64	0.59	0.31
4	0.80	0.67	0.44	37	0.66	0.59	0.33	70	0.70	0.62	0.34
5	1.86	1.05	0.47	38	1.86	1.05	0.47	71	0.64	0.59	0.33
6	0.70	0.62	0.36	39	0.64	0.59	0.36	72	0.66	0.59	0.29
7	1.6	0.78	0.37	40	0.70	0.59	0.37	73	0.66	0.59	0.29
8	1.86	1.05	0.37	41	0.64	0.59	0.39	74	0.70	0.59	0.30
9	0.80	0.78	0.34	42	0.66	0.58	0.33	75	0.66	0.58	0.28
10	0.70	0.62	0.32	43	0.80	0.62	0.36	76	0.64	0.58	0.28
11	0.64	0.59	0.47	44	0.66	0.58	0.29	77	0.66	0.59	0.27
12	0.66	0.58	0.41	45	0.64	0.59	0.29	78	0.80	0.62	0.36
13	0.80	0.67	0.47	46	0.64	0.59	0.28	79	0.66	0.59	0.28
14	0.66	0.58	0.37	47	0.64	0.59	0.28	80	0.64	0.59	0.28
15	0.64	0.59	0.39	48	0.66	0.62	0.29	81	0.70	0.59	0.29
16	0.70	0.62	0.37	49	0.70	0.62	0.29	82	0.70	0.59	0.29
17	0.66	0.62	0.36	50	0.66	0.62	0.29	83	1.6	0.67	0.47
18	1.86	1.05	0.39	51	1.6	1.05	0.72	84	0.70	0.62	0.36
19	0.64	0.59	0.32	52	0.80	0.67	1.01	85	0.66	0.62	0.39
20	0.70	0.62	0.29	53	0.70	0.27	0.27	86	1.6	0.78	0.72
21	0.64	0.58	0.30	54	0.80	0.78	0.52	87	0.80	0.78	0.72
22	0.64	0.59	0.52	55	1.86	0.78	0.44	88	0.70	0.59	0.28
23	0.66	0.59	0.44	56	0.70	0.67	0.36	89	0.66	0.59	0.28
24	0.70	0.62	0.47	57	0.70	0.59	0.31	90	1.86	1.05	0.72
25	0.66	0.58	0.37	58	0.80	0.78	1.05	91	0.64	0.58	0.27
26	0.70	0.62	0.39	59	0.70	0.59	0.33	92	0.64	0.59	0.28
27	0.70	0.59	0.36	60	0.66	0.59	0.31	93	0.66	0.62	0.29
28	1.6	0.78	0.39	61	0.70	0.67	0.39	94	0.70	0.67	0.31
29	0.70	0.62	0.33	62	1.6	0.78	0.72	95	0.66	0.59	0.33
30	0.66	0.59	0.30	63	1.86	1.05	1.01	96	0.70	0.67	0.36
31	0.70	0.62	0.30	64	1.6	1.05	1.83	97	0.70	0.62	0.39
32	1.6	1.05	0.72	65	0.70	0.67	0.47	98	1.6	1.05	1.83
33	1.86	1.05	1.01	66	0.70	0.67	0.44	99	0.70	0.78	0.72
								100	0.70	0.79	0.72

يلاحظ من جدول (17) أن "توافيق" التي ظهرت للأخطاء المعيارية في تقدير القدرة كانت بأربع حالات مثلت النسب كما يلي:

(%95) = $S_3 < S_2 < S_1$

(%2) = $S_1 < S_3 < S_2$

(%2) = $S_2 < S_1 < S_3$

(%1) = $S_1 < S_2 < S_3$

ويظهر هنا أن الدقة في القياس أعلى في الاختبار الثالث(جميع الفقرات) يليها الاختبار التكيفي يليها الاختبار الاستطلاعي وأن أكبر نسبة للخطأ المعياري في تقدير القدرة ظهرت عند استخدام المرحلة الأولى من الاختبار (مرحلة الاختبار الاستطلاعي) .

رابعا : النتائج المتعلقة بالسؤال الرابع :

ما مقدار معامل الارتباط بين قيمة (θ) في كل مرحلة مـن مراحـل القيـاس ومحـك العلامات المدرسية ؟

استخدمت علامات الطلبة المدرسية التي حصلوا عليها من خلال ما أعده المعلمون من اختبارات يومية وشهرية لوحدتي الضرب والقسمة للكشف عن فاعلية القياس التكيفي حيث تـم إيجاد معامل ارتباط بيرسون بين قيم (θ) والعلامات المدرسية والجدول (18) يبين قيم معاملات الارتباط بين العلامات المدرسية وقيم (θ).

الجدول (18) معاملات الارتباط بين العلامات المدرسية وقيم (θ)

العلامات المدرسية	θ1	θ2	θ3
	0.74	0.79	0.80

يتبين من الجدول (18) أن معاملات الارتباط متقاربة لكن أعلاها كان بين تقدير القدرة باستخدام جميع الفقرات والعلامات المدرسية، يليه معامل الارتباط بين قـيم القـدرة باسـتخدام مرحلة القياس التكيفية الثانية والعلامات المدرسية، في حين كان معامل الارتباط بين قيم القدرة في مرحلة الاختبار الاستطلاعي والعلامات المدرسية أدناها.

الفصل الخامس
مناقشة النتائج والتوصيات

الفصل الخامس
مناقشة النتائج

يتناول هذا الفصل مناقشة للنتائج التي توصلت إليها الدراسة لبناء اختبار تكيفي لقياس القدرة الرياضية وفق الاستراتيجية ثنائية المرحلة في نظرية الاستجابة للفقرة وكان من ابرز النتائج التي تم التوصل إليها في هذه الدراسة:

بعد حذف الأفراد غير المطابقين للنموذج وحذف الفقرات غير المطابقة للنموذج (نموذج المعلم الواحد) تراوحت قيم الصعوبة للفقرات بين (1.12و2.88-) لوجيت وتراوح الخطأ المعياري في تقدير معلم صعوبة الفقرات بين (0.11 و 0.17) لوجيت. ولتحديد القدرة الأولية للمفحوصين استخدم الاختبار الاستطلاعي لتوجيه المفحوصين إلى اختبار المرحلة الثانية من القياس التكيفي وقد تراوحت قدرات المفحوصين في هذه المرحلة بين (3.55إلى 0.88-) لوجيت بمتوسط مقداره (0.74) لوجيت وتراوح الخطأ المعياري في تقدير قدرة المفحوصين بين (1.86إلى 0.7) لوجيت.

تم استخدام الاختبار التكيفي الثنائي المرحلة بعد الاختبار الاستطلاعي وقد تراوحت قدرات المفحوصين بعد تحليل البيانات بين (3.70 و 0.34-) لوجيت وتراوح الخطأ المعياري بين (1.85و 0.59) لوجيت بمتوسط للقدرة بلغ "(0.82) لوجيت ومتوسط للخطأ المعياري بلغ (0.68) لوجيت.

وبعد ذلك تم تطبيق جميع فقرات الاختبار على عينة القياس التكيفي جميعهم (100) طالب وطالبة وبعد تحليل البيانات تراوحت قدراتهم بين (3.53و0.44-) لوجيت وتراوح الخطأ المعياري في قياس هذه القدرة بين (1.83و0.28) لوجيت بمتوسط للقدرة بلغ (1.85) لوجيت ومتوسط للخطأ المعياري بلغ (0.42) لوجيت.

بعد الإطلاع على القيم الناتجة نلاحظ أن قياس القدرة من خلال فقرات الاختبار التكيفي جميعها أعطى اكبر قيمة لمتوسط القدرة أما عند قياس القدرة من خلال فقرات الاختبار الاستطلاعي فحصلنا على اقل قيمة لمتوسط القدرة وربما يعود السبب في ذلك هو صغر عُمر الطلبة عينة الدراسة (الصف الثاني الأساسي) وعدم ألفتهم بمثل هذه الاختبارات (اختبار الاختيار من بين البدائل) وقد اتفقت هذه النتيجة مع ما توصل إليه نور الدين (2002) في دراسته التي أجراها حيث لاحظ أن الاختبار التكيفي يعطي دقة أعلى من الاختبار الاستطلاعي وذلك لأن الفقرات في الاختبار التكيفي تنتقى حيث تناسب مع قدرة الفرد وحافظت القدرة المقاسة بواسطة فقرات الاختبار التكيفي الثنائي على مرتبة الوسط بين القدرتين السابقتين. وبذلك تعارضت هذه النتيجة مع ما توصل إليه هندركسون (2002) ونور الدين في إن القياس التكيفي أكثر دقة في تقدير قدرة المفحوصين حيث كانت النتيجة في هذه الدراسة أن فقرات الاختبار التكيفي الثنائي حافظت على مرتبة الوسط بين القدرتين المقاستين بواسطة فقرات الاختبار الاستطلاعي وجميع فقرات الاختبار.

وعند ملاحظة نسبة الخطأ المعياري في قياس القدرات نلاحظ أن متوسط قيمة الخطأ عند استخدام الاختبار الاستطلاعي أعطت اكبر متوسط للخطأ، أما عند استخدام الاختبار التكيفي الثنائي فكان متوسط قيمة الخطأ في قياس القدرة يقل بشكل ضئيل عن متوسط خطا القياس عند استخدام الاختبار الاستطلاعي فكانت تقل بمقدار (0.06) لوجيت وإذا قارنا بين متوسط القدرة باستخدام الاختبار الاستطلاعي والاختبار التكيفي الثنائي نجد أن الفرق ضئيل أيضاً (0.82) لوجيت متوسط تقدير القدرة بواسطة الاختبار التكيفي الثنائي مطروح منه (0.74) لوجيت متوسط تقدير القدرة بواسطة الاختبار الاستطلاعي والناتج = (0.08) لوجيت وهو فرق بسيط

بين القدرتين وان دل ذلك فيدل على فعالية الاختبار الاستطلاعي في تقدير قدرات عينة القياس التكيفي وهم طلبة الصف الثاني الأساسي قليلوا الخبرة في الامتحانات والحنكة في اختيار البديل الصحيح من بين البدائل المعروضة وهذا يدعم الهدف من هذه الدراسة وهو بناء اختبار تكيفي لقياس القدرة الرياضية وفق الاستراتيجية ثنائية المرحلة في نظرية الاستجابة للفقرة لقياس قدرة طلبة الصف الثاني الأساسي وللكشف عن ملاءمة هذا النوع من الاختبارات لطلبة المرحلة الأساسية الأولى من سلم التعليم. وقد اتفقت هذه النتيجة مع دراسة الثوابية (2004) والبرصان (2006) في أن نسبة الخطأ المعياري قي تقدير القدرة باستخدام جميع فقرات الاختيار أعطت أقل قيمة للخطأ المعياري بين القيم الناتجة عند استخدام الاختبار الاستطلاعي أو الاختبار التكيفي الثنائي في قياس القدرة.

وعند مقارنة القدرة والخطأ في قياس هذه القدرة بين مراحل القياس التكيفي جميعها نجد أن:

- الاختبار الذي أعطى أكبر متوسط للقدرة كان الأقل في متوسط الخطأ.
- الاختبار الذي أعطى اقل متوسط للقدرة أعطى اكبر متوسط للخطأ في قياس تلك القدرة.
- حافظ الاختبار التكيفي الثنائي على الحد الأوسط لكل من متوسط القدرة ومتوسط الخطأ بين الاختبارين السابقين. وهذا يدل على ملاءمة الاختبارات التكيفية للطلبة صغار السن (طلاب الصف الثاني الأساسي).

وقد اتفقت نتائج هذه الدراسة مع ما توصل إليه نور الدين (2002) ولورد (1971) في دراستيهما. فبواسطة الاختبار التكيفي يمكن قياس قدرات الطلبة بعدد قليل من الفقرات ففي هذه الدراسة استخدمت الباحثة (12) فقرة في كل اختبار تكيفي لتقيس قدرات الطلبة التي وزعتها أيضا على

(5) مستويات وقد أعطت النتائج تقديرات أعلى مما أعطته عنـد اسـتخدام فقـرات الاختبـار الاستطلاعي بالرغم من حصول الباحثة على تقديرات أعلى من ذلك عندما اسـتخدمت جميـع فقرات الاختبار ألـ(60) على نفس العينة التي خضعت للاختبارات الفرعية (12) فقـرة وقـد عزت السبب في ذلك هو عدم الجدية في الإجابات من قبل الطلبة أثنـاء تعرضـهم للاختبـار التكيفي (12) فقرة أو قد تكون عدم الألفة بهذا النوع من الاختبارات أو الخوف.

إن التباين في تقدير قدرة المفحوصين يعتمد كما أشار لندن وجـلاس (Linden and Glas,2002) على مقدار التطابق بين صعوبة فقرات الاختبار وقـدرات المفحوصـين ففـي النموذج أحادي المعلمة يقدم الاختبار اكبر كمية من المعلومات عندما تتطابق صعوبة الفقـرات مع قدرات الطلبة. فيلاحظ بأن نتائج الدراسة قد أشارت إلى أن قيم تقدير (θ) قد تباينـت عبـر مراحل القياس الثلاث: مرحلة القيـاس الاسـتطلاعي، مرحلـة القيـاس باسـتخدام الاختبـار الاستطلاعي واختبار المرحلة الثانية، ومرحلة القياس باستخدام جميع فقرات الاختبار، فقد كـان أعلى متوسط لتقدير القدرة باستخدام جميع فقرات الاختبار التكيفي وقد بلغ (1.85) لوجيت أمـا أدنى متوسط تقدير قدرة المفحوصين فقد ظهر في مرحلة تقدير القدرة باستخدام فقرات الاختبار الاستطلاعي حيث بلغ متوسط التقدير (0.74) لوجيت وكانت قيمة متوسط التقدير في مرحلـة الاختبار التكيفي الثنائي هي المرحلة بين التقديرات (0.82) لوجيت.

ولمقارنة نتائج هذه الدراسة بمحكات خارجية تم إيجاد معاملات الارتباط بين قـيم (θ) وعلامات الطلبة المدرسية وقد تراوحت قيم معاملات الارتباط بين قيم (θ) والعلامات المدرسية لاختبارات الرياضيات في وحدتي (الضرب والقسمة) التي أعدها المعلمـون مـا بـين (0.80 و0.74) حيث كان أعلى معامل ارتباط بين العلامات المدرسية وقيم القدرة التي تم حسابها عند استخدام جميع فقرات الاختبار يلي ذلك معامل الارتباط بين العلامة المدرسية وقيم القدرة التـي

تم الحصول عليها من تطبيق الاختبار التكيفي الثنائي وكان اقل معامل ارتباط بين العلامات المدرسية وقيم القدرة عند استخدام الاختبار الاستطلاعي حيث بلغ (0.74).

تشير هذه النتائج بشكل عام إلى فاعلية القياس التكيفي في تقدير القدرة للمفحوصين من مرحلة التعليم الأساسي الأولى حيث تقاربت قيم معاملات الارتباط بين قيم (θ) عبر مراحل القياس المختلفة والعلامات المدرسية كمحكات خارجية كما أن تقدير القدرة بواسطة الاختبار التكيفي الثنائي أعطى المفحوص تقديراً دقيقاً لقدرته بالاعتماد على الفرق الضئيل عند استخدم الاختبار الاستطلاعي وكان الوسط بين تقدير القدرة بالاختبار الاستطلاعي واختبار تكيفي مؤلف من جميع فقرات الاختبار كذلك حافظ على الوسط في الخطأ المعياري بين الاختبارين السابقين.

ومن المتوقع أن يتقارب الخطأ المعياري بدرجة اكبر لو تم أجراء مرحلة ثالثة من القياس التكيفي كذلك يلاحظ بأن القياس التكيفي قد عمل على تقليل الوقت اللازم لتطبيق الاختبار وذلك من خلال اختصار عدد الفقرات من (60) فقرة إلى (12) فقرة. وبهذه النتيجة اتفقت مع ما توصل إليه شيرمز وستمر (1996) في دراستهما.

التوصيات

ومن الجدير بالذكر أن هذه الدراسة لم تتحقق من فاعلية القياس التكيفي عند استخدام وحدات مختلفة لنفس المبحث لذلك توصي الباحثة بإجراء المزيد من دراسات القياس التكيفي والتي تتناول الاختبارات التحصيلية للمبحث كاملاً كذلك توصي الباحثة بإجراء اختبارات تكيفية لمباحث أخرى لنفس المرحلة (مرحلة التعليم الأساسي الأولى) كذلك أجراء دراسات للقياس التكيفي باستخدام الحاسوب وضرورة أن تعمل الجامعات على توفير البرامج اللازمة لأغراض البحث العلمي مثل توفير البرمجيات الأزمة للقياس التكيفي المحوسب.

المراجع

المراجع العربية

البرصان، إسماعيل سلامة. (2006). **أثر عدد مراحل القياس التكيفي وعدد أسئلة كل مرحلة في تقدير القدرة والخطأ المعياري في تقدير القدرة والخطأ المعياري فـــي التقــدير باستخدام أسئلة ثنائية ومتعددة التدريج**. رسالة الدكتوراه غير منشورة. جامعة عمان للدراسات العليا، عمان، الأردن.

الثوابية، أحمد محمود. (2004). **فاعلية القياس التكيفي في تقويم التحصيل في مبحث الأحياء للصف الثاني الثانوي العلمي في الأردن**. رسالة دكتوراه غير منشورة. جامعة عمان للدراسات العليا، عمان، الأردن.

الشريفين، نضال كمال .(2003) . **مدى تحقق معايير الفاعلية في معادلة اختبارين إحـــداهما ثنائي التدريج والأخر متعدد التدريج وفق نماذج النظرية الكلاسيكية والنظرية الحديثة في القياس**. رسالة دكتوراه غير منشورة. جامعة عمان العربيـــة للدراســـات العليـــا. عمان، الأردن.

عبد العزيز، نبيل عبدالله. (1994). **تطوير ومعيرة مقياس القدرة الرياضية للفئة العمرية من 13-15 سنة في الأردن**، رسالة ماجستير غير منشورة. جامعـــة اليرمـــوك. أربـــد. الأردن.

عبد العزيز، نبيل عبدالله. (2006). **فاعلية طرق منحنى خصائص الفقرة في الكشف عن الأداء التفاضلي لمتغير الجنس في فقرات اختبار رياضيات متعددة التدريج وثنائية التدريج.** رسالة دكتوراه غير منشورة. جامعة عمان العربية للدراسات العليا. عمان. الأردن .

العجيلي، صباح.(2000). **استخدام الاستراتيجية ثنائية المرحلة في بناء اختبار محبـوك فـي قواعد اللغة العربية للصف السادس الابتدائي.** مجلة الأستاذ. 21(1),47-86.

علام، صلاح الدين. (1995). **الاختبارات التشخيصية مرجعية المحك في المجالات التربويـة والنفسية والتدريبية.** القاهرة. دار الفكر العربي.

العموش، جميل ضيف الله. (2003). **فاعلية القياس التكيفي في تقويم بعض القدرات المعرفية لدى طلبة السنة الأولى الجامعية.** رسالة دكتوراه غير منشورة. جامعة عمان العربية للدراسات العليا. عمان. الأردن.

الفرجات، هشام عقلة. (2004). **بناء بنك أسئلة لمبحث الكيمياء للصف الثاني الثانوي العلمي.** رسالة ماجستير غير منشورة. جامعة مؤتة. الكرك. الأردن.

كاظم، أمينة. (1998). **استخدام نموذج راش في بناء اختبار تحصيلي في علم النفس وتحقيق التفسير الموضوعي للنتائج.** مطبوعات جامعة الكويت.

نور الدين، أمين .(2002). **فاعلية استخدام الاختبار الموائم باستخدام الحاسوب في تقدير قدرة الأفراد وتحديد الخصائص السيكومترية للمقياس.** رسالة دكتـوراه غيـر منـشورة. جامعة عين شمس.

المراجع باللغة الإنجليزية

Alkhadher, O. and Clarke, D. (1998). Equivalence and Predictive validity of paper and computerized adaptive formats of the differential aptitude tests. *Journal of Occupational and Organizational Psychology*. 71 (3), 205-218.

Allen, M.J and yen, W.M. (1979). *Introduction to measurement Theory. Californian: Cole Publishing company*.

Anastasi, A. urbina, S. (1997): *Psychological Testing. (7th ed.)*, New York: Printice Hall.

Baker, F. (2001). *The Basics of Item Response theory*. Maryland university of Maryland.

Brengstrom, B. A. Lunz, M.E. Gershon, R.C (1992): Altering the level of difficulty in computer adaptive testing, *Applied Measurement in Education,* 5 (2), 137-149.

Burghof, K.I (2001). *Assembling an item –bank for computerized linear and adoptive testing in Geography. Inter national Education*. 2(4) retrieved September 10, 2007 for http://www.flindersedu.au/eduaction/ieJ.

Castle R. (1997). *Relative efficiency of tst vers TMCT using IRT*, Linclon, Nerbras ka.

Chislli, E. Campbell, J. & Zedecks. (1981): *Measurement theory for the behavioral science san Francisco: freeman company*.

Crist, S. (1989). *Computerized adaptive tests: Eric Pigest No. 107*, Eric. Database http://ericanet/eda/ed315425.htm.

Crocker, L. and Algina, J. (1986) *Introduction to classical and Modern test theory*. New your: Holt Pine hart and Winston.

De beer, M. and Visser, D (1998). *Compare ability of paper & Pencil & computerized adaptive versions of the general scholastic aptitude Test (GAST) senior: South African Journal of Psychology*, Vol. 28. Issue. 1.

Dunkel, P. (1999). *Consideration in Developing and using computer adoptive tests to assess second language proficiency Eric Digits*, EDO-FL, 99-07.

Emberston, S.E. Reise, S. P. (2000). *Item Responses theory for psychologists*. Mahwah, NewJersy. Lawrence Erlbaum Associates, Publishers.

Haladyna, M. Royd. G. H. (1983). A comparison of Two Approaches to Criterion Referenced Test construction. *Journal of Educational Measurement* 20 (3) 271-282.

Hambleton, R. swaminathan, H. and Rogers, H. (1991). *Fundamental of Item response theory*. Newyork: sage publication the international professional Publishers.

Hambleton, R.K. and Swami Nathan, H. (1985). *Item Response theory principle and Application*. Kluwer: Nijhoff Publishing.

Hattie, J. (1985*). Methodology Review Assessing Unidimensionality of test and item*. Applied Psychological Measurement. 99: 139-164.

Hend rickson, A. (2002). *Scaling of Two stage Adaptive Test configurations for Achievement Testing*. Unpublished Doctoral Dissertation. University of Iowa.

Hulin, C. L. Drasgow. F. & Parwsons. C.K. (1983). *Item response theory: Application to psychological measurement*. Illinois: Dou-Jones – Irwin.

Julian, E. R. (1988). *Using computerized Patient simulations to measure the clinical competence of physicians*. Applied Measurement in Education.

Kim, H. Plake, B (1993): *Monte Carlo simulation comparison of two – stage testing and computerized Adaptive testing*, ERIC: ED 357041.

Koch W. Dod, B (1989). *An Investing of Procedures for communized adoptive testing using partial credit scoring Applied measurement in Education.*

Larkin, K. C. Weis, D, J. (1975). A*n Empirical investigation of computer administrated middle Ability Testing ERIC: ED* 096343.

Linarce, J.M., & wright, B.D. (1993). *Auser Guide to BIGSTEPS (Computer Program)*. Version 2.4, Chicago: MESA press.

Lincare, J. (2000). *Computer- Adoptive Testing methodology whose Time Has com*. Seol: Komeza Press.

Linden, W. and Glas. C. (2002*). Computerized Adoptive testing theory and practice: Lituwer Academic Publishers*.

Lord, F. (1980). *Application of Item Response Theory to Practical Testing problems*. New Jersey: Lawrence Erl Baum Associates publishers.

Lord. (1971). *Theoretical study of two stage testing Psychometrical*, 36:227-242

Schinkpke, D. and Green, F. (1995). A comparison of Item Selection Routines in Linear and Adoptive tests, *Journal of Educational measurement* 32 (3), 227-242.

Shermis, M & Stemmer, P. (1996). *Computerized Adoptive skills Assessment in state wide Testing: Research on computing in Education*. Vol.29:49.

Shermis, M. & Ful Keuson, J. (1996) Computerized adoptive moth tests for Elementary Talen Development selection. Reopen Review. 19 (2), 91-96.

Visser, M. (1998) Adaptive testing comparison. South African university. *Journal of psychology*, 63, (vol).28.

Warm, A. (1978). *Aprimer of Item Response Theory: US. Coast Guard Institute Oklahoma* 73/69.

Weiss, D.J. (1983). *New horizons in testing: latent trit test theory and Computerized adoptive testing*. MN: Assessment systems corporations.

الملاحق

ملحق (1) أهداف عامة للوحدتين
أهداف التدريس لوحدة القسمة
الفصل الثاني

يتوقع من الطالب: بعد دراسته لوحدة القسمة أن يكون قادراً على أن:

1. يوضح عملية القسمة على أنها توزيع بالتساوي
2. يذكر حقائق القسمة على أنها طرح متكرر
3. يذكر حقائق القسمة على الأعداد ضمن حقائق الضرب
4. يحل مسائل حياتية باستخدام العمليات على الأعداد (خطوة واحدة)

أهداف التدريس لوحدة الضرب

يتوقع من الطالب بعد دراسته لوحدة الضرب أن يكون قادراً على أن:

1- يمثل عملية الضرب على أنها جمع متكرر باستخدام أشياء محسوسة

2- يذكر حقائق للضرب في العدد2

3- يذكر حقائق للضرب في العدد3

4- يذكر حقائق للضرب في العدد4

5- يذكر حقائق للضرب في العدد5

6- يذكر حقائق للضرب في العدد10

7- يذكر حقائق للضرب في العدد1 وفي العدد صفر

8- أن يحل مسائل مستخدماً العمليات على الأعداد خطوة واحدة

ملحق (2) فقرات الاختبار بالصورة الأولية بعد التحكيم
اختبار رياضيات في وحدتي (الضرب والقسمة) للصف الثاني الأساسي
زمن الاختبار: حصتان

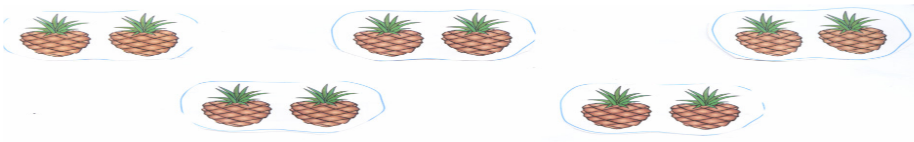
اسم الطالب:

المدرسة:

تعليمات الاختبار:

- يتكون هذا الاختبار من (64) فقرة جميعها من نوع الاختيار من متعدد ولكل فقرة ثلاث بدائل-
- المطلوب منك الإجابة عن كل فقرة من فقرات الاختبار وذلك باختيار البديل الصحيح من بين البدائل الثلاث المعروضة ووضع دائرة حوله-
- لا تختار أكثر من بديل واحد لكل سؤال حتى لا تضيع عليك العلامة المخصصة لذلك السؤال-

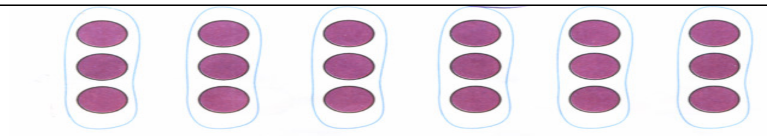
1



جملة الضرب التي تعبر عن عدد حبات الأناناس في (5) مجموعات في كل منها (2) حبة أناناس هي:

أ- 5×2= 10 ب- 2+2+2+2+2= 10 ج- 3×5= 15

2



عدد الدوائر كحاصل ضرب =

أ. 3+3+3+3+3+3=18 ب. 6×3=18 ج. ست مجموعات في كل منها ثلاث دوائر

3



عدد العصافير عن طريق جملة الجمع المكرر=

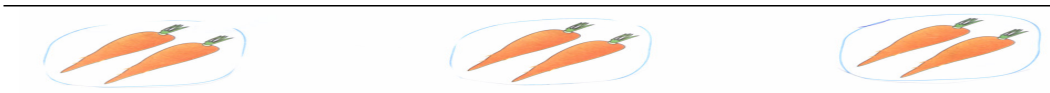
أ- 4×4=16 ب- 8+8=16 ج- 4+4+4=16

4

عندما تضرب عدد في العدد (1) يكون الناتج هو؟

أ- نفس العدد المضروب ب- العدد مكرر ج- صفر

5



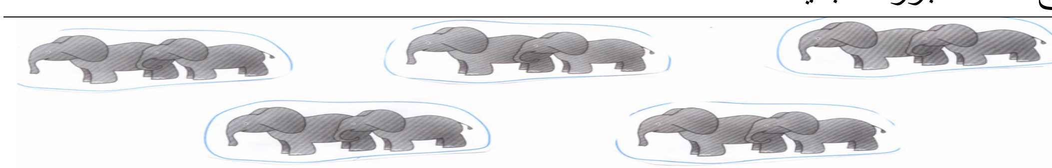
3×2=6 هي جملة ضرب معبرة عن الصورة السابقة ويدل العدد 2 على:

أ- عدد المجموعات.

ب- عدد حبات الجزء في كل مجموعة.

ج- عدد الجزرات جميعاً.

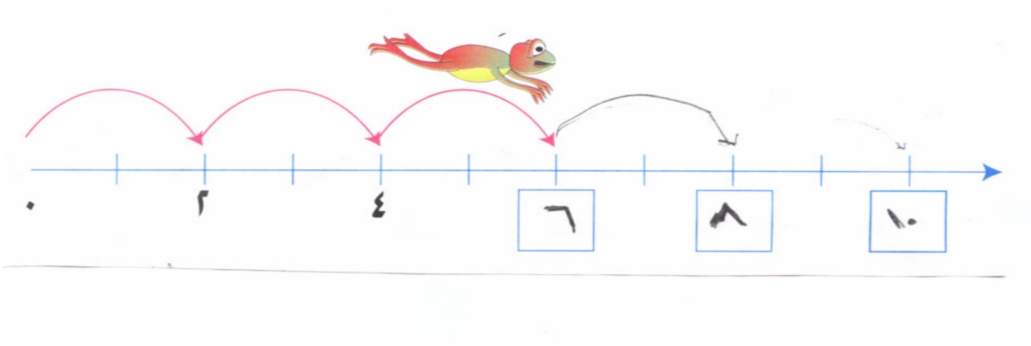
6



جملة الضرب المعبرة عن الصورة السابقة هي؟

أ- 2×5= 10 ب- 2+2+2+2+2=10 ج- عشر فيلة

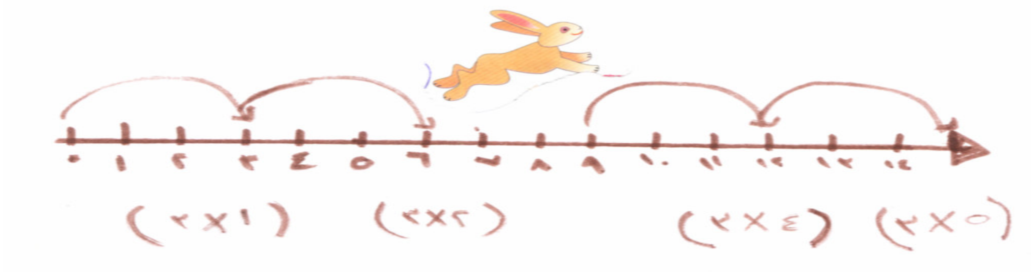
7



حسب الرسم السابق في الأعلى فإن الضفدع يصل إلى موقعه بعد قفزة بطول؟

أ- 2 وحدة　　　ب- 6 وحدة　　　ج- صفر وحدة

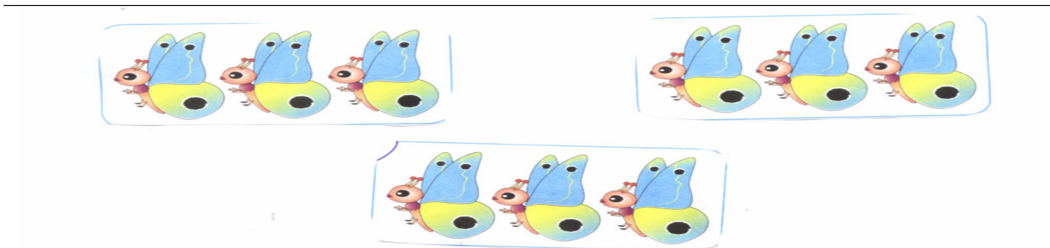
8



القفزة المفقودة للأرنب حسب الشكل السابق هي؟

أ- 3×2　　　ب- 3×3　　　ج- 3×1

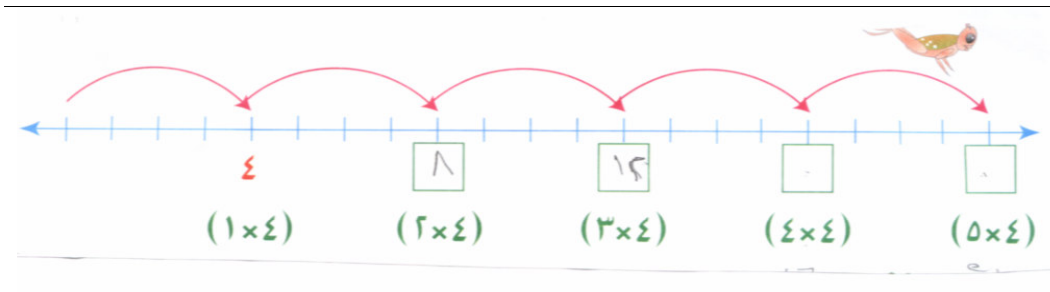
9



جملة الضرب المعبرة عن عدد الكل في الصورة السابقة هي؟

أ- 9×0= 9　　　ب- 9×1=9　　　ج- 3×3=9

10



العددان اللذان سيقف عليهما الجندب بعد كل قفزة هما؟

أ- (16) (20)　　　ب- (12) (14)　　　ج- (16) (18)

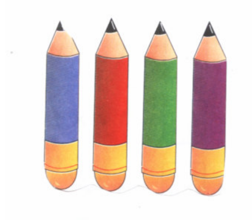
11 اشترك في المسابقات الثقافية والدينية (5) طلاب من كل صف من الصفوف الأول والثاني والثالث والرابع، واستأجرت المدرسة حافلة لنقل الطلاب، فإذا كانت الحافلة تتسع لاثنين وعشرين راكباً فهل تتسع لنقل الطلاب؟ علماً بأنه سيرافقهم معلم واحد فقط

أ- ستتسع الحافلة لنقل الطلاب

ب- لن تتسع الحافلة لنقل الطلاب

ج- ستتسع الحالة للمتسابقين من طلبة الصف الأول والثاني فقط

12



- إذا أردت شراء (4) أقلام رصاص وكان ثمن القلم الواحد (5) قروش فكم قرشاً ستدفعه للبائع؟

أ- ثلاثون قرش　　ب- عشرون قرش　　ج- خمسة وعشرون قرش

13



جملة الضرب المعبرة عن الصورة السابقة هي؟

أ- 40=10×4　　ب- 400=100×4　　ج- 4000=1000×4

---

14 =10×5

أ- 40　　ب- 50　　ج- 10

---

15 عندما تضرب عدداً في العدد (صفر) يكون الناتج هو؟

أ- صفر　　ب- نفس العدد المضروب　　ج- العدد المكرر

16



اشترى عادل (3) بناطيل من البناطيل المعروضة في الصورة فكم ديناراً دفع؟

أ- (13) دينار　　ب- عشرة دنانير　　ج- 30 دينار

---

17 أن العددان اللذان إذا ضربناهما يكون ناتجهما = 12 هما؟

أ- 3×3 | ب- 3×4 | ج- 12×0

18 أن تمثيل ضرب 3×5 كجملة جمع هو؟

أ- 5+5+5 | ب- 3+5+5+2 | ج- 15

19


اشترت المدرسة (3) كرات من الكرات المعروضة في الصورة السابقة فما ثمنها؟

أ- 8 دنانير | ب- 6 دنانير | ج- 3 دنانير

20 **أربع خمسات هي جملة ضرب يعبر عنها رياضياً كما يلي؟**

أ- 4×5=20 | ب- 20 | ج- 4×5

21 **إذا أردت شراء (3) علب عصير وكان ثمن العلبة الواحدة (7) قروش فكم ستدفع للبائع؟**

أ- 14 قرش | ب- 9 قروش | ج- 21 قرش

22 (3) سلات من التفاح في كل منها (5) تفاحات، فإن عدد التفاحات جميعها =

أ- 15 تفاحة | ب- 22 تفاحة | ج- 51 تفاحة

23 4×3 تساوي

أ- اثنا عشر | ب- أربعة عشر | ج- عشرون

24 2×4 يساوي

أ- 4×2 | ب- 9 | ج- 4×4

25 عدد النجمات في (4) مجموعات إذا كان عددها في كل مجموعة نجمتان=

أ- 8 | ب- 4 | ج- 9

26 5×4 يساوي؟

أ- 4 | ب- 5 | ج- 20

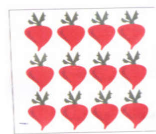
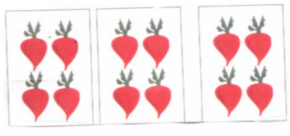
27	5×1 يساوي؟	أ- 1	ب- 4	ج- 5
28	اشترت تهاني (3) أقلام وكان ثمن القلم الواحد (3) قروش فكم قرشاً ستدفع تهاني للبائع؟	أ- 9	ب- 6	ج- 10
29	3×4 يساوي؟	أ- 12	ب- 9	ج- 16
30	31- إذا أراد أحمد بيع القلم الواحد بـــ (2) قرش فكم يكون ثمن قلم واحد؟	أ- 2	أ- 1	أ- 4
31	2×2 يساوي؟	أ- 2	ب- 4	ج- 6
32	$0\times5=$	أ- صفر	ب- 5	ج- 1
33	$7\times0=$	أ- صفر	ب. 1	ج. 7
34	$(2\times4)+(1\times2)=$	أ- 8	ب- 10	ج- 2
35	$(9\times1)-(1\times3)=$	أ- 6	ب- 16	ج- 26
36	$15\times1=$	أ- صفر	ب- 1	ج- 15

37

25 ×0 =

أ- 25 ب- 15 ج- صفر

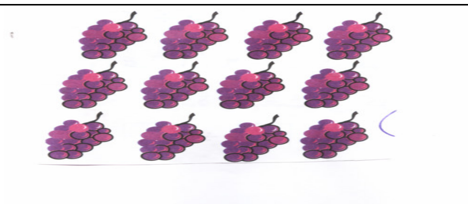
38

اشترى والدك حزمة فجل فيها (12) حبة فجل وزعتها والدتك على (3) أكياس فكم فجلة تضع في الكيس الواحد؟

أ- 3 فجلات ب- 4 فجلات ج- فجلتان

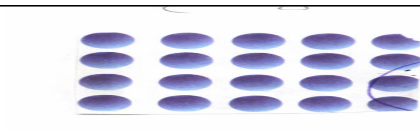
39



قسم رائد (12) قطف عنب على (4) صحون، فكم قطف وضع في الصحن الواحد؟

أ- 3 ب- 4 ج- 2

40



لو حصرت كل (5) دوائر معاً فكم مجموعة يصبح لديك؟

أ- 5 ب- 4 ج- 2

41

8÷2= 4 مسألة المقسوم فيها هو.......

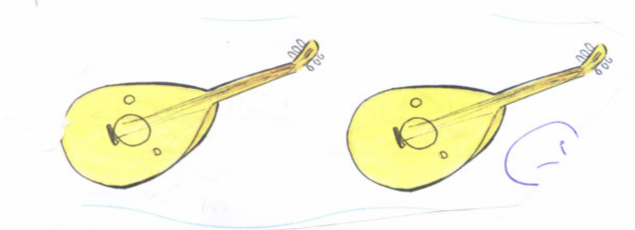
أ- 4 ب- 2 ج- 8

42

كم مرة تطرح (3) من (12) ليكون الباقي صفر؟

أ- 3 ب- 2 ج- 4

43



كم (2) يوجد في العدد (2)؟

أ- 2 　　 ب- 1 　　 ج- صفر

44

4÷20 = ☐

أ- 5 　　 ب- 4 　　 ج- 2

45

1÷5 = ☐

أ- 1 　　 ب- 2 　　 ج- 5

46



48- ذهب (30) طالب في رحلة مدرسية إلى شاطئ العقبة وقرروا ركوب القوارب فإذا كان كل قارب يتسع لثلاث طلاب فقط فكم قارب نحتاج لركوب الطلاب جميعاً؟

أ- 10 قوارب　　ب- 8 قوارب　　ج- 30 قارب

47

العدد الذي إذا ضربته بـــ (5) كان الناتج (20) هو:

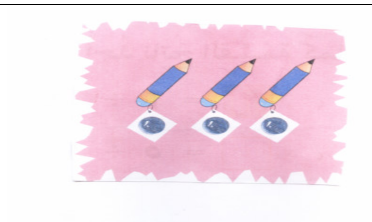
أ- 10　　ب- 4　　ج- 20

48

عندما تقسم عدداً على العدد (1) فإن الناتج ؟

أ- نصف العدد المقسوم　　ب- صفر　　ج- نفس العدد المقسوم

49



اشترت فرح (3) أقلام بثمن (30) قرش فكم ثمن القلم الواحد؟

أ- عشر قروش　　ب- ثلاثة قروش　　ج- سبعة قروش

50

أعطاك والدك (20) دينار لتوزعها على (4) عائلات فقيرة بالتساوي فكم تعطي لكل عائلة؟

أ- 3 دنانير　　ب- 5 دنانير　　ج- 4 دنانير

51

اختر العملية المناسبة　21 ☐ 7 = 3

أ- +　　ب- ÷　　ج- ×

52	53- العملية المناسبة للمسألة التالية: 20 ☐ 5 = 4 أ- + ب- × ج- ÷		
53	قامت وزارة الزراعة بزراعة (50) شجرة في خمس شوارع بالتساوي فكم شجرة زرعت في كل شارع؟ أ- 10 شجرات ب- 4 شجرات ج- 5 شجرات		
54	- العدد المقسوم في مسألة القسمة هو العدد؟ أ- الأول في الترتيب ب- الثاني في الترتيب ج- الثالث في الترتيب		
55	(6) تفاحات وزعت على (3) طلاب من طلبة الصف الثاني تمثيل هذه المسألة هو؟ أ- 3×2=6 ب- 6÷3=2 ج- 6-صفر=6		
56	 قسمت الأم (3) قطع حلوى على أبنائها الثلاثة بالتساوي، فكم قطعة يأخذ كل واحد منهم؟ أ- 3 قطع ب- قطعة ونصف ج- قطعة واحدة		
57	9÷3 عملية قسمة ناتجها هو ؟ أ- 2 ب- 3 ج- 9		

58	ثلاثة تقسيم ثلاث = واحد تمثيل هذه المسألة رياضياً هو؟		
	أ- 3÷3=1	ب- 1÷3=	ج- 3÷3=3
59	التوزيع المتساوي" يعني استخدام عملية؟		
	أ- الضرب	ب- القسمة	ج- الجمع
60	كم مرة تطرح (5) من (15) ليصبح الباقي صفر؟		
	أ- 3 مرات	ب- 4 مرات	ج- ولا مرة
61	8÷4=2 مسألة المقسوم عليه فيها هو؟		
	أ- 8	ب- 4	ج- 2
62	عملية القسمة عملية تبني على عملية هامة جداً تسبقها وعليك إتقانها ..... وهي عملية ....؟		
	أ- الجمع	ب- الطرح	ج- القسمة
63	(3×3) ÷3 =		
	أ- 9	ب- 3	ج- 6
64	10 ÷ 2 =		
	أ- 10	ب- 20	ج- 5

ملحق (3) نموذج تقويم الاختبار

الرقم	الفقرات	موافق بشدة	موافق	غير موافق	غير موافق بشدة
-1	تعليمات الاختبار واضحة ومحددة				
-2	فقرات الاختبار مصاغة بلغة سليمة				
-3	فقرات الاختبار خالية من الأخطاء العلمية				
-4	فقرات الاختبار خالية من الأخطاء الطباعية				
-5	فقرات الاختبار تغطي أهداف الوحدة كافة				
-6	فقرات الاختبار تغطي المهارات الأساسية لمجالات التقويم				
-7	فقرات الاختبار ترتبط ارتباطاً كبيراً بالأهداف				
-8	فقرات الاختبار موزعة على موضوعات الوحدة بشكل مناسب				
-9	المطلوب من الفقرات واضح ومحدد				
-10	بدائل الفقرات مموهات جيدة				
-11	زمن الاختبار ملائم				
-12	تنسيق وإخراج الاختبار مناسب				

الفقرات التي تعتقد أنها بحاجة إلى تعديل

رقم الفقرة	سبب التعديل	التعديل المقترح

ملحق(4) أسماء المحكمين

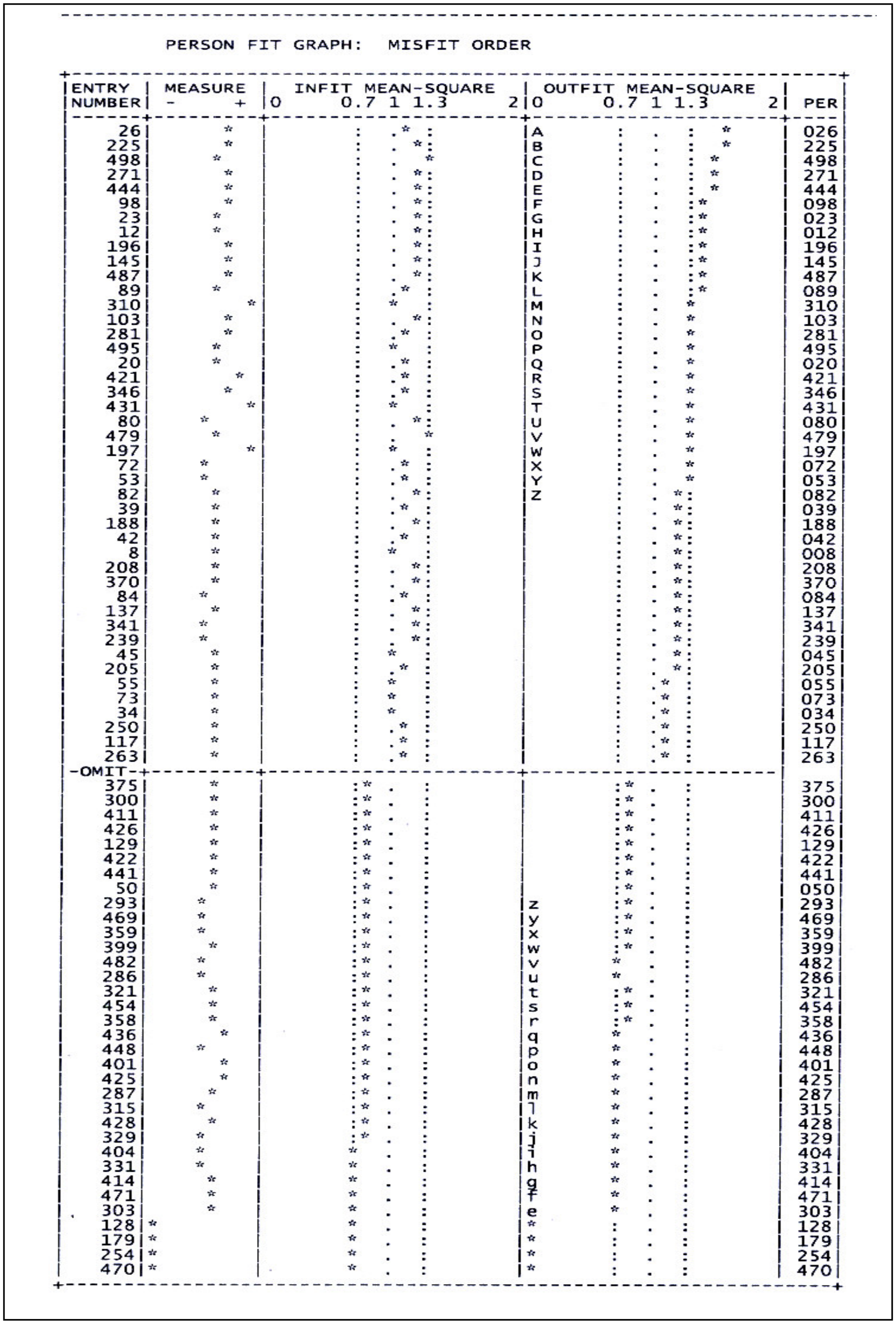
رائدة معرعر	معلمة رياضيات / مدرسة السعادة الأساسية المختلفة
عهد عظمات	معلمة صف / مدرسة الثلاج الأساسية
ابتسام شلهوب	معلمة تربية / مدرسة البوادي الأساسية
فاتن المساعيد	معلمة صف / مدرسة البشرية الأساسية
نسرين الشرعة	معلمة صف / مدرسة روضة بسمة الأساسية
د. عبد الحافظ الشايب	مدرس قياس وتقويم في جامعة آل البيت
سمارة عظمات	مشرف تربوي / تربية البادية الشمالية الشرقية
مفلح الرحال	رئيس قسم الإشراف / مديرية تربية البادية الشمالية الشرقية
د. مروان الصمادي	مدرس في جامعة العلوم والتكنولوجيا
د. ثائر غباري	مدرس في جامعة الزرقاء الأهلية
هالة حداد	مديرة مدرسة متقاعدة
د. حسين مراشدة	مدرس في جامعة اليرموك
د. سعيد الطيطي	مشرف رياضيات / مديرية تربية البادية الشمالية الشرقية
د. محمد الاطرش	مدرس في جامعة الزرقاء الأهلية

ملحق (5) إحصائيات المطابقة الداخلية والخارجية لجميع الأفراد

PERSON STATISTICS: MISFIT ORDER

ENTRY NUMBER	RAW SCORE	COUNT	MEASURE	ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTBIS CORR.	PER
26	44	64	0.85	0.28	1.16	1.4	1.69	3.7	-0.12	026
225	51	64	1.46	0.32	1.25	1.3	1.64	2.2	-0.39	225
498	33	64	0.08	0.26	1.34	5.0	1.56	5.4	-0.36	498
271	50	64	1.36	0.31	1.26	1.4	1.53	2.0	-0.35	271
444	51	64	1.46	0.32	1.23	1.2	1.52	1.8	-0.32	444
98	51	64	1.46	0.32	1.22	1.1	1.48	1.7	-0.28	098
23	30	64	-0.12	0.26	1.30	4.3	1.48	4.9	-0.28	023
12	30	64	-0.12	0.26	1.28	4.1	1.44	4.5	-0.24	012
196	50	64	1.36	0.31	1.21	1.2	1.43	1.7	-0.26	196
145	50	64	1.36	0.31	1.22	1.2	1.40	1.6	-0.26	145
487	50	64	1.36	0.31	1.22	1.2	1.40	1.6	-0.26	487
89	35	64	0.22	0.26	1.16	2.3	1.40	3.8	-0.05	089
310	58	64	2.39	0.43	1.06	0.2	1.40	0.8	-0.06	310
103	44	64	0.85	0.28	1.29	2.5	1.40	2.3	-0.33	103
281	51	64	1.46	0.32	1.15	0.8	1.40	1.4	-0.16	281
495	40	64	0.56	0.27	1.02	0.3	1.38	2.8	P0.16	495
20	36	64	0.28	0.26	1.11	1.6	1.36	3.3	Q0.04	020
421	54	64	1.79	0.35	1.11	0.5	1.35	1.0	-0.10	421
346	51	64	1.46	0.32	1.14	0.8	1.34	1.3	-0.12	346
431	59	64	2.60	0.47	1.04	0.1	1.34	0.6	-0.02	431
80	21	64	-0.76	0.28	1.20	1.8	1.33	2.3	-0.11	080
479	35	64	0.22	0.26	1.31	4.5	1.33	3.2	-0.30	479
197	60	64	2.84	0.52	1.05	0.1	1.33	0.5	-0.06	197
72	27	64	-0.32	0.26	1.16	2.1	1.30	3.0	-0.04	072
53	18	64	-0.99	0.29	1.19	1.3	1.30	1.7	-0.07	053
82	30	64	-0.12	0.26	1.28	4.0	1.29	3.1	-0.23	082
39	30	64	-0.12	0.26	1.13	1.9	1.28	3.0	0.02	039
188	35	64	0.22	0.26	1.27	3.9	1.27	2.6	-0.22	188
42	28	64	-0.26	0.26	1.13	1.8	1.26	2.7	0.03	042
8	34	64	0.15	0.26	1.04	0.7	1.26	2.6	0.15	008
208	32	64	0.01	0.26	1.25	3.8	1.26	2.8	-0.19	208
370	32	64	0.01	0.26	1.25	3.8	1.26	2.8	-0.19	370
84	25	64	-0.46	0.27	1.14	1.7	1.25	2.3	0.01	084
137	35	64	0.22	0.26	1.25	3.6	1.24	2.4	-0.18	137
341	27	64	-0.32	0.26	1.24	3.0	1.24	2.4	-0.16	341
239	27	64	-0.32	0.26	1.23	3.0	1.24	2.4	-0.15	239
45	28	64	-0.26	0.26	1.08	1.2	1.22	2.3	0.10	045
205	39	64	0.49	0.26	1.20	2.4	1.22	1.8	-0.11	205
55	30	64	-0.12	0.26	1.05	0.7	1.20	2.2	0.16	055
73	30	64	-0.12	0.26	1.05	0.7	1.20	2.2	0.16	073
34	32	64	0.01	0.26	1.02	0.3	1.20	2.2	0.21	034
250	31	64	-0.05	0.26	1.19	2.9	1.19	2.1	-0.08	250
117	29	64	-0.19	0.26	1.16	2.4	1.17	1.8	-0.03	117
263	35	64	0.22	0.26	1.14	2.2	1.14	1.4	0.00	263
	BETTER	FITTING	OMITTED							
375	31	64	-0.05	0.26	0.88	-2.1	0.85	-1.8	0.47	375
300	32	64	0.01	0.26	0.87	-2.3	0.85	-1.8	0.49	300
411	32	64	0.01	0.26	0.87	-2.3	0.85	-1.8	0.49	411
426	31	64	-0.05	0.26	0.86	-2.3	0.84	-1.9	0.49	426
129	38	64	0.42	0.26	0.86	-2.1	0.83	-1.7	0.50	129
422	33	64	0.08	0.26	0.85	-2.5	0.84	-2.0	0.50	422
441	33	64	0.08	0.26	0.85	-2.5	0.84	-2.0	0.50	441
50	38	64	0.42	0.26	0.85	-2.1	0.82	-1.7	0.50	050
293	25	64	-0.46	0.27	0.85	-1.9	0.82	-1.9	z0.52	293
469	27	64	-0.32	0.26	0.84	-2.3	0.81	-2.1	y0.54	469
359	25	64	-0.46	0.27	0.84	-2.0	0.82	-1.9	x0.54	359
399	28	64	-0.26	0.26	0.84	-2.4	0.81	-2.2	w0.54	399
482	19	64	-0.91	0.28	0.84	-1.4	0.77	-1.7	v0.55	482
286	24	64	-0.53	0.27	0.84	-2.0	0.80	-2.0	u0.55	286
321	35	64	0.22	0.26	0.84	-2.7	0.81	-2.1	t0.53	321
454	35	64	0.22	0.26	0.84	-2.7	0.81	-2.1	s0.54	454
358	31	64	-0.05	0.26	0.83	-2.9	0.81	-2.3	r0.55	358
436	44	64	0.85	0.28	0.83	-1.7	0.78	-1.5	q0.53	436
448	20	64	-0.83	0.28	0.82	-1.6	0.76	-1.9	p0.57	448
401	48	64	1.18	0.30	0.82	-1.3	0.73	-1.4	o0.54	401
425	48	64	1.18	0.30	0.82	-1.4	0.70	-1.6	n0.56	425
287	31	64	-0.05	0.26	0.82	-3.2	0.79	-2.6	m0.58	287
315	20	64	-0.83	0.28	0.81	-1.7	0.75	-2.0	l0.59	315
428	31	64	-0.05	0.26	0.81	-3.2	0.79	-2.6	k0.58	428
329	26	64	-0.39	0.26	0.81	-2.7	0.77	-2.6	j0.60	329
404	25	64	-0.46	0.27	0.80	-2.6	0.76	-2.5	i0.62	404
331	20	64	-0.83	0.28	0.79	-2.0	0.72	-2.2	h0.63	331
414	39	64	0.49	0.26	0.79	-3.0	0.75	-2.4	g0.62	414
471	39	64	0.49	0.26	0.78	-3.1	0.74	-2.4	f0.63	471
303	38	64	0.42	0.26	0.78	-3.3	0.75	-2.5	e0.63	303
128	1	64	-4.38	1.02	0.74	-0.3	0.09	-1.4	d0.48	128
179	1	64	-4.38	1.02	0.74	-0.3	0.09	-1.4	c0.48	179
254	1	64	-4.38	1.02	0.74	-0.3	0.09	-1.4	b0.48	254
470	1	64	-4.38	1.02	0.74	-0.3	0.09	-1.4	a0.48	470
MEAN	31.	64.	-0.05	0.30	1.00	-0.1	1.00	0.0		
S.D.	14.	0.	1.13	0.08	0.10	1.2	0.18	1.2		

ملحق 6 رسم توضيحي للمطابقة (جميع الأفراد)



ملحق (7) إحصائيات المطابقة الداخلية والخارجية لجميع الفقرات

ITEMS STATISTICS: MISFIT ORDER

ENTRY NUMBER	RAW SCORE	COUNT	MEASURE	ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTBIS CORR.	ITEM
1	323	477	-0.96	0.11	1.26	5.6	1.60	6.0	A0.07	I001
5	244	477	-0.12	0.10	1.28	6.8	1.43	6.4	B0.13	I005
2	285	477	-0.55	0.10	1.20	4.9	1.40	5.3	C0.17	I002
32	220	477	0.13	0.10	1.32	7.3	1.36	5.3	D0.11	I032
3	314	477	-0.86	0.11	1.17	4.0	1.32	3.6	E0.17	I003
38	292	477	-0.62	0.10	1.17	4.3	1.24	3.2	F0.20	I038
6	263	477	-0.32	0.10	1.22	5.5	1.23	3.5	G0.18	I006
41	215	477	0.18	0.10	1.21	5.0	1.22	3.3	H0.20	I041
55	197	477	0.37	0.10	1.18	4.0	1.21	3.1	I0.25	I055
63	132	477	1.12	0.11	1.20	3.3	1.21	2.1	J0.22	I063
14	238	477	-0.06	0.10	0.99	-0.3	1.19	3.1	K0.41	I014
37	165	477	0.72	0.11	1.17	3.3	1.17	2.1	L0.24	I037
22	287	477	-0.57	0.10	0.97	-0.7	1.17	2.3	M0.39	I022
8	264	477	-0.33	0.10	1.12	3.1	1.15	2.3	N0.27	I008
17	225	477	0.07	0.10	1.11	2.8	1.13	2.1	O0.30	I017
33	208	477	0.25	0.10	1.02	0.5	1.12	1.9	P0.39	I033
29	437	477	-2.84	0.17	1.11	0.8	0.88	-0.5	Q0.20	I029
9	259	477	-0.28	0.10	1.09	2.5	1.06	0.9	R0.31	I009
11	212	477	0.21	0.10	1.06	1.4	1.08	1.4	S0.35	I011
64	172	477	0.64	0.11	1.03	0.7	1.07	1.0	T0.37	I064
7	226	477	0.06	0.10	1.07	1.8	1.03	0.5	U0.34	I007
13	323	477	-0.96	0.11	0.96	-1.0	1.07	0.8	V0.37	I013
24	284	477	-0.54	0.10	1.05	1.3	1.06	0.9	W0.33	I024
18	200	477	0.34	0.10	1.06	1.4	1.06	0.9	X0.35	I018
21	198	477	0.36	0.10	0.89	-2.6	1.06	0.9	Y0.50	I021
16	264	477	-0.33	0.10	1.04	1.0	1.05	0.9	Z0.36	I016
	BETTER	FITTING	OMITTED							
30	259	477	-0.28	0.10	0.96	-1.2	0.89	-1.8	z0.44	I030
60	157	477	0.81	0.11	0.95	-1.0	0.92	-1.0	y0.44	I060
51	182	477	0.53	0.10	0.93	-1.7	0.95	-0.8	x0.48	I051
40	155	477	0.84	0.11	0.94	-1.2	0.87	-1.7	w0.47	I040
10	224	477	0.08	0.10	0.94	-1.5	0.92	-1.4	v0.46	I010
19	298	477	-0.69	0.10	0.94	-1.5	0.87	-1.9	u0.44	I019
36	276	477	-0.45	0.10	0.93	-2.0	0.94	-1.0	t0.45	I036
42	217	477	0.16	0.10	0.94	-1.6	0.89	-1.9	s0.46	I042
35	203	477	0.30	0.10	0.92	-2.0	0.93	-1.2	r0.49	I035
26	261	477	-0.30	0.10	0.92	-2.3	0.88	-2.0	q0.48	I026
45	268	477	-0.37	0.10	0.91	-2.3	0.85	-2.4	p0.48	I045
61	149	477	0.91	0.11	0.91	-1.8	0.84	-2.0	o0.48	I061
49	201	477	0.33	0.10	0.90	-2.4	0.83	-2.8	n0.50	I049
54	160	477	0.78	0.11	0.90	-2.1	0.83	-2.3	m0.50	I054
12	281	477	-0.51	0.10	0.90	-2.9	0.89	-1.7	l0.48	I012
47	203	477	0.30	0.10	0.89	-2.8	0.84	-2.7	k0.51	I047
31	284	477	-0.54	0.10	0.88	-3.3	0.81	-3.1	j0.50	I031
59	168	477	0.69	0.11	0.86	-3.1	0.83	-2.3	i0.53	I059
58	188	477	0.47	0.10	0.85	-3.6	0.80	-3.2	h0.55	I058
43	190	477	0.44	0.10	0.83	-4.0	0.78	-3.6	g0.57	I043
48	198	477	0.36	0.10	0.83	-4.3	0.80	-3.4	f0.57	I048
46	165	477	0.72	0.11	0.82	-4.0	0.76	-3.5	e0.57	I046
23	278	477	-0.47	0.10	0.80	-5.8	0.72	-4.7	d0.58	I023
57	184	477	0.51	0.10	0.80	-4.9	0.73	-4.3	c0.60	I057
53	203	477	0.30	0.10	0.79	-5.4	0.77	-3.9	b0.60	I053
56	222	477	0.10	0.10	0.78	-6.1	0.72	-5.2	a0.62	I056
MEAN	232.	477.	0.00	0.10	1.00	-0.2	1.01	-0.1		
S.D.	55.	0.	0.63	0.01	0.13	3.1	0.18	2.6		

Abstract

Shdeifat, Sabah Jamil. A construction of an adaptive testing to measure students mathematical ability according to the second stage strategy in Item response theory. A. thesis, Yarmouk University, 2008 (Supervisor: Dr. Nedal Kamal Al-Sharifeen).

This study aims at constructing an adaptive testing to measure students mathematical ability, according to the second stage strategy in Item Response Theory. To achieve this purpose; the scholan constructed multiple choice test for second graders in mathematics; units (division, multiplication), these two units consisted of 60 items, in total.

The test included three samples:

1. Primitive testing sample of 100 students.
2. Choice gradation sample of 500 students.
3. adaptive measure sample of 100 students.

These sample were chosen form primary schools in Mafraq's three direction of Education.

This test was applied to the study sample, and results were analyzed using software's: (Notepad, BIGSTEPS, SPSS). The attained evaluation of individuals ability, Item difficulty values, and standers errors, measured with the suitability of statistics to the form of the test, resulted the efficacy of adaptive measure in measuring the abilities of young student in the stage of primary education (Second graders).

These results showed a high level accuracy in estimating individuals abilities while using the using the adaptive testing instead of the traditional testing.

Key words: Paragraph response theory, second stage strategy, adaptive testing, individual's ability, paragraph difficulty.